School of Electronic Engineering and Computer Science

The Neural Detection of Emotion In Naturalistic Settings.

A thesis submitted to Queen Mary, University of London in partial fulfilment of the requirements for the degree of Master of Philosophy.
March 2015

Author: Tajinder Singh Dhami

Primary Supervisor: Professor Patrick Healey
Secondary Supervisor: Professor Peter McOwan

Declaration.

I, Tajinder Singh Dhami, confirm that the research included within this thesis is my own work or that where it has been carried out in collaboration with, or supported by others, that this is duly acknowledged below and my contribution indicated. Previously published material is also acknowledged below.

I attest that I have exercised reasonable care to ensure that the work is original, and does not to the best of my knowledge break any UK law, infringe any third party's copyright or other Intellectual Property Right, or contain any confidential material.

I accept that the College has the right to use plagiarism detection software to check the electronic version of the thesis.

I confirm that this thesis has not been previously submitted for the award of a degree by this or any other university.

Signature: [ ]

Date: 8th March 2015

Abstract

The Field of Emotion research has experienced resurgence partially due to the interest in Affective Computing, which includes calls for natural emotion to be studied in natural type settings. A new generation of commercial mobile EEG headsets present the potential for new forms of experimental design that may move beyond laboratory settings. Across the Arts and Cultural sectors there are longstanding questions of how we may objectively evaluate creative output, and also subjective responses to such artefacts.

This research adjoins these concerns to ask; How can low-cost, portable EEG devices impact on our understanding of cultural experiences in the wild?

Using a commercial emotiv Epoch EEG headset, we investigated gauging Valence and Arousal levels across the two contrasting experimental settings of a live theatre performance, and a controlled laboratory setting.

Our results found that only Valence could be reliably detected, and only with a good degree of confidence in laboratory settings. This determines that we may only be able to gather very general information regarding cultural experiences via the enlisted EEG technology and methods, and only in controlled conditions.

Acknowledgements.

Firstly, I would like to thank my supervisors both current and prior, Professor Patrick Healey, Professor Peter McOwan, and Dr Hatice Gunes, for all their patient advice, guidance, support and assistance.

I would like to thank all the institutions and individuals who allowed permission to access the necessary public and private spaces to make recordings, and for providing support in conducting the pilot studies and formal experiments, namely; The Serpentine Gallery, Tate Modern, The Hayward Gallery, The Lisson Gallery, The Pace Gallery, The Barbican, The Soho Theatre, and The Bush Theatre.

I would like to thank Holger, and Kosta for explaining some basic signal processing techniques, and also the Matlab 12012b interface.

Further I would like to extend my warm gratitude to The Research Office, Richard Kelly, Professor Clive Parini & the wider QMUL services that assisted in times of need, and made me feel a valuable part of the student community.

Importantly I would also like to thank all my MAT peers, who share in this exciting and challenging journey of creative's researching in an Electrical Engineering & Computer Science environment, in particular cohorts (9,10,11). I would also like to thank the EPSRC for funding this research as part of the MAT Doctoral Training program.

Finally I would like to thank all my Friends and Family, and especially my darling Anoushka who makes me beam from ear to ear.

Contents

List of Figures.

List of Tables.

CHAPTER 1

Introduction.

Figure 1.1: An iconic promotional image of the film Brainstorm 1983, Directed by Douglas Trumbull[1]

1.1 Introduction

Douglas Trumbull's 1983 classic science fiction work Brainstorm (see Fig 1) features the development of a non-invasive neural headset that records subjects perceptual sensory experiences onto magnetic tape. Replaying these encodings allows others to sensually re-experience these perceptions. Unexpectedly, the developers discover that the headset simultaneously encodes felt emotions and emotionally tinged thoughts as inherent processes within perception.

The title of this research project is, The Neural Detection of Emotion in Naturalistic Settings. This project critically investigates the central question; How can low-cost, portable EEG devices impact on our understanding of cultural experiences in the wild? Thus, in this research we use the emotiv Epoch EEG headset to gauge neural

[1] Retrieved from : http://www.engadget.com/2004/10/29/movie-gadget-friday-the-brain-scanner-from-brainstorm/

[2] Retrieved from http://www.audiostream.com/content/download-week-masaki-batoh

emotional responses to cultural productions via Arousal and Valence levels. Through conducting two main studies in contrasting experimental settings, we demonstrate the important differences between lab-based, and 'in-the-wild' settings. For the 'in-the-wild' setting a live theatre performance was enlisted, whilst for the laboratory setting emotional video-clips were used.

At the outset of this project it was hoped that through our experiments we might be able to determine the viability of the enlisted portable EEG headset to present opportunities for objective insight into subjective cultural experiences. Two speculative idealistic applications were envisaged for a successful process, which in real world scenarios may serve the creative sector. Firstly, such a process may serve creators in the construction, presentation, and evaluation of works. Secondly, such a process may allow the emotions to become a new form of creative material in creative endeavours.

Through our investigations we may be able to determine a process to meet our intentions and potentially and lay a foundation for a fuller framework in future research.

1.2 Neural Detection: Brain Imaging Technology.

In recent decades a range of non-invasive Neuro-imaging technologies have emerged that allow unprecedented levels of observable access to the working human brain. Previously, such observational opportunities were reliant upon chance occurrences of injury, post-mortem or the use of proxy species. These new imaging technologies allow for the visualization of both brain structure and function.

The most prominent structural mapping technologies are Computed Tomography (CT) and Magnetic Resonance Imaging (MRI). Both take the form of a large-scale scanner, into which an individual is horizontally placed. The CT scanner captures an array of Omni-directional x-rays that are combined together to create single 2d and 3d images. The MRI scanner uses a magnetic field and pulsed radio waves to align and then dislodge the protons of hydrogen atoms. The measured movements and alignments of the atoms can be computed into highly detailed 3d images that differentiate between different tissues and the structures of the brain. Despite their effectiveness, both of these scanners are immobile due to their scale, weight, and operational speciality. Further their high costs ensure they are only accessible at specific locations.

The MRI scanner is also used for functional observation in a technique called functional Magnetic Resonance Imaging (fMRI). Here, the fMRI measures changes in Blood-Oxygen-Level-Dependent (BOLD) contrasts, by tracking changes in blood flow to detect the transitions in magnetisation between oxygen-rich and oxygen-poor blood as its basic measure. Whilst this results in very noisy data, the underlying signal can be extracted statistically and computed to produce high quality imaging revealing which parts of the brain are active or being activated at any given time. Thus, this is s a valuable technique for researchers and scientists wanting to investigate task related neural responses.

The Positron Emission Tomography (PET) scanner, uses low dosages of a radioactive material known as a tracer, which is injected into an individuals bloodstream and then travels to collects in their tissue and organs. A ring of detectors are used to detect pairs of gamma rays emitted by the tracer at any given time in the area being observed. As blood flows through the brain in response to a task or stimulus, the active areas produce higher quantities of gamma ray pairings, making them highly visible when the data is computed into images. The PET scanner also provides the measure of the degree of activity at any site. As with the previously mentioned scanners the PET scanner is a large scale, specialist, and immobile technology.

Magnetoencephalography (MEG) is another non-invasive functional imaging technique. It measures the brains magnetic fields as generated by neuronal activity, and calculates their spatial distribution to produce high quality 3d images. It has excellent spatial and temporal resolution. However, its measuring of magnetic fields requires it to be placed within a special shielded room, to prevent its readings being obscured by other signals.

A more mobile and inexpensive technology for obtaining information of neural activity is the Electroencephalograph (EEG). EEG is a technique for detecting brain activity as measured by voltage fluctuations on the scalp. However, the presence of the scalp between the point of brain activity and the EEG electrode creates a blurring of the signal, which results in a low spatial resolution. Thus single neuron firing cannot be detected; rather it is the simultaneous firing of millions of neighbouring neurons in unison. Also as EEG can only detect activity in the cortical regions it is unable to provide any detailed topographic data. The EEG has the advantage of a high temporal resolution (in the order of milliseconds), which makes it ideal for detecting continuous neural responses and changes in activity. Two prominent EEG detection techniques have developed to account for both continuous and discrete timeframes. Event-related Potentials (ERP's) measure voltages in minute discrete time-locked responses to any given task. The responses can either be averaged over multiple trails, or taken as a single trail for analysis. Spectral Power (SP) is a continuous method that traditionally categorizes the SP of measured voltages into frequency bandwidths labelled Alpha, Beta, Delta, Theta and Gamma. Each is associated with a key brain state such as; deep sleep, light meditation, relaxed, waking consciousness, or high level processing.

EEG experiments have traditionally been located in the laboratory type settings. This due both to its sensitivity to artefacts incurred by movement and its sensor set-up, which are wired to signal amplifiers and other recording peripherals. A new generation of low cost, portable, off-the-shelf commercial headsets have began to emerge. These offer to potentially remove the constraints of the EEG laboratory set up, and some of the issues associated with it; uncomfortable sensors, lengthy set up times, and wiring. The new mobile headsets promise new neural detection contexts and forms of experimental design. The headsets are increasing being successfully used in formal academic studies for example; Liu and Sourina (2012), Liu, Sourina and Hafiyyandi (2013), Badcock et. al (2013), Debener et. al (2013). These are validating their usage and confidence as formal research tools. (A detailed exploration of EEG and its usage in emotion detection are provided in chapter 3)

1.3 Arts Evaluation : Subject & Object.

In the Arts, questions of how we may make objective evaluations of cultural and creative output are longstanding. Stretching back to Immanuel Kant's (1892, trans Bernard 1951) Critique of Judgment, questions arose about subjective-objective aesthetic relationships when considering Art forms. Firstly, Kant's proposal was to differentiate between the conditions of beauty and sublimity, and then to imply two further conditions within sublimity of 'The mathematical' and 'The dynamical' conditions. His thesis implied forms of objective measurability to his 'Mathematical' definition, and pure subjectivity in regards to his notion of the 'Dynamic'.

One approach towards objective evaluation has been based on statistics such as attendance figures, linger times, questionnaires, and economic value. However these may in turn be linked to marketing strategies, accessibility, cultural capital rather than the 'creative thing' in question. Whilst these may be justifiable within a particular remit they provide no heuristic or enhancive value to the creative.

Projects such as Neuroaesthetic's aim to make ground into unlocking the nature of Aesthetic brain responses, yet largely focus on the neural correlates of viewing (beauty), and as such serve as forms of neural topography and mapping (Zeki, 1999).

Formal frameworks for the objective evaluation of cultural products have also been proposed. Goldman (2004) proposed the notion of the ‘Ideal Critic’ operating in ‘Ideal Conditions’. Here the critic would have an extensive knowledge of the contextual envelopes of the work being assessed, and have ideal attributes of undivided attention where objectivity reigned, and further be sensitive to golden insights of the works affective nature and potential influence on a broad spectrum of viewers. This view expanded beyond classical frameworks of valuing beauty and representation to cater for modern and contemporary art forms that often function to oppose, satirize, extend and distort their foundational conventions. Any resulting Ideal critic’s assessment would provide a benchmark against which all other evaluations could be graded.

Frameworks have also emerged in Psychology, which extend creative output to include the fields of science, math, business and organisation. Sternberg & Lubert's (1996) **c**onfluence model contains components of knowledge, thinking styles, personality, motivation and environment. Here, creativity is constructed of multiple attributes which operate in at least three different forms; processes, domains and styles (Sternberg, 2005).

Social psychologist Teresa M Amabile's (1996, 2012) presents a similar perspective in her Componential theory of creativity. Her grounding premise for creativity is the production of ideas or outcomes towards a goal demonstrative of both novelty and appropriateness. Her model is comprised of 3 components; Expertise (Domain relevant Skills), Creative skills (Creativity- relevant process), and Motivation (Intrinsic task motivation), all based within a social environment of production. Her interlinked modules function within a dual evaluation construct of consensual and conceptual definition. The consensual relates to; appropriate observers who are knowledgeable or familiar with a products domain of production, who all agree that it is creative. This may also be applied to its process of construction. The conceptual relates to; any output will be seen as creative if it should demonstrate qualities of appropriate novelty to the domain, a useful, correct, or valuable response to the task at hand, and that such output be deemed heuristic as opposed to algorithmic.

This progression from prior social psychological formulas of assessing creativity through forms of associative thought responses to specific stimuli, is shared by Sandra Russ (1993), who points to the role of affect and play in the creative process. As an evaluation structure Russ considers output over process, highlighting qualities of uniqueness, originality, novelty, adaptability, and aesthetic pleasantness according to the standard of a particular discipline.

A commonality in all these frameworks is a focus on a distanced evaluation inclusive of context as a fulfilment of objective assessment. These do not necessarily focus on the experience it generates. In new forms and presentations of media and arts this notion of the distance between the audience and a creative object is increasingly diffused. In interactive works, the audience may become active participants leading to modular or generative encompassing experiences. Jaak Panksepp's (1988, 2000, 2010) theories of emotion indicate the presence of priori behaviours before survivalist motivations where there is a sense of simple playfulness and exploration. As these new modes of creative practise extend within digital technologies further original considerations are being made as to how these experiences may be assessed in terms of their effectiveness.

Ernest Edmonds & Brigid Costello (2007) have developed a framework within which such interactive works can objectively be methodologically designed and assessed for affectivity. Their principle is to designate interactive experiences as playful behaviour, from which they extract the term 'Pleasure' as a prime component of play, citing its aspects of joy, delight and amusement. They have produced a 13-category survey method alongside a video-recall interview process that they believe allows the discernment of the works success.

Thus we may sense a movement towards contemporary objective evaluations that include considering the subjective experiences of a viewer. Such processes are invaluable in formulating new forms of evaluation, and to contribute towards accounting for a wider diaspora who engage with cultural output, over an 'expert' who may have their own motivations.

1.4 Neural Signals as a Creative Material.

Figure 1.2: Examples of seminal Artworks that use EEG signals to drive their content. (Left), Lucier's 'Music for a Solo Performance' 1965 [2], and Mori's Wave UFO, 2003 [3] (right).

In Arts practise, Alvin Lucier's seminal 1965 work, 'Music for a solo performance' was the first artwork to directly use brain signals. Lucier attached two EEG electrodes to his scalp and used the shifting Alpha frequencies to trigger and resonate musical instruments to create a cacophonous orchestral sonic composition (see Fig 1.2). Works which followed in this tradition include Richard Teitel Bame 1968, Nam June Paik's, A Tribute to John Cage 1973, Christoph De Boecks Staalhammer 2003, and Luciana Hail's Psi Chic 2007. In all these works the dynamism of the Alpha frequency served as a form of narrative, relating to transitional states of relaxation and excitation.

This approach has also been adopted by audio-visual artists, for example David Rosenboon's, Ecology of the Skin 1971, and extended to incorporate varying brainwave states of Theta, Alpha, and Beta, in Mariko Mori's seminal work 'Wave UFO' 2003. In this work that attempts to represent the Buddhist principles of inter-subject-connectivity associated with the deeper aspects of consciousness, three visitors enter a UFO shaped vessel (see Fig 1.2). The changes in their brains states alongside mental activity and facial artefacts detected via EEG, transform orb shaped forms projected onto the ships ceiling

---

[2] Retrieved from http://www.audiostream.com/content/download-week-masaki-batoh
[3] Retrieved from http://www.deitch.com/artists/sub.php?artistId=15

Figure 1.3. Examples of Contemporary Interactive Artworks which utilise EEG signals; (Top left), Smart studio's competitive mediation game, Brain Ball [4], Interaxon's Levitation chair controlled by Alpha spectral power levels [5] (Right), and Moura, Guimaraes, and Canibal's, live performance work 'Camara Neuronal' 2000 [6] (Bottom left)

Contemporary Interactive works sharing this intuitive theme include Smart Studio's Brain Ball, 1999. Here two audience members compete in a meditative gaming situation where the goal is to move a ball into the opponents half (see Fig 1.3). The users levels of meditation and calmness as computed from their EEG's ball controls the ball, and the more relaxed signals move the ball. Canadian interaction firm Interaxon adopt a similar idea in their Levitation Chair (see Fig 1.3). By increasing their alpha levels through meditation a user causes the chair to ascend. This is also echoed in XXXY's infinity Simulator 2012, where a user can experience the sensation of flying. Here, neural Alpha signals are routed to a trapeze rig and the user can arise and descend in the environment based on their levels of alpha activity.

---

[4] Retrieved from http://diccan.com/Images/brainball_press.jpg
[5] Retrieved from http://scienceline.org/2012/01/mind-over-matter/
[6] Retrieved from http://jmartinho.net/camara-neuronal/

João Martinho Moura, Miguel Pedro Guimaraes, and Adolfo Lúxuria Canibal's, live performance collaboration 'Camara Neuronal' 2000 is another example of how creatives are extending the traditions of mediums in their use of EEG signals (see Fig 1.3) .

Artist Eva Lee and Affective Neuroscientist James Coan attempt to build a bridge between the Sciences and the Arts in their work 'Discrete Terrain' 2007. This is a data visualisation of EEG data harvested in controlled conditions. Here, the signal oscillations and their classifications are causal of the changes in the visual elements that are reminiscent of bar charts rendered as an abstract terrain.

Thus in these few highlights it may be visible that having access to the hidden spaces of the brain is an area of great interest for creatives. Whilst many of the above works, use simple levels of excitation and calmness to make works, having a protocol for accessing the emotions may potentially allow for rich detailed art forms to develop where the emotions and emotion data streams may be used as a form of creative material.

1.5 Thesis Layout.

The content of this thesis is arranged as thus; Chapter 2 presents a foundational literature review of the field of Emotion Research. It highlights developments of the four main frameworks in the field; Discrete, Dimensional, Appraisal and Cognitive. Chapter 3 unveils the viability and support for the neural detection of emotion via EEG. Chapter 4 highlights issues surrounding the call for emotion studies to move beyond laboratory conditions in to real world settings. Included in this chapter are details of a number of pilot test's to find the most appropriate setting for this research project. Chapter 5 is a transparent rendition of the signal processing techniques used in this project. Chapter's 6 and 7 detail the experiments conducted in two settings; natural (Theatre) and controlled (Laboratory) conditions. Finally chapter 8 is the conclusive text, which surmises this projects findings.

CHAPTER 2

Emotion Theory

2.1: Introduction

In his publication 'Emotional Intelligence', Daniel Goleman presents an intriguing passage of text. This may be read as suggesting that the emotions are a central process in the evolution of the biological, the neurological and in turn conscious experience; "The most primitive root of our emotional life is the sense of smell...is it edible, should I repel it, can I mate with it" (Goleman, 1996). Thus emotion research may be seen as investigating the complexity of phenomena that is holistically entwined within a body and its historical and temporal interactions within the space of its surrounds.

As a first movement, this body of research looks to map the wide terrain of emotion research to the present day to gain an understanding of the field. Whilst the emotions register in Philosophy under the term 'The Passions', formal objective scientific inquiry into these phenomena began over 100 years ago. In the duration to our present day a vast wealth of ideas, propositions and findings have been published. It would be a near impossible task to attempt to give a detailed account of these here, thus in this chapter a basic overview highlighting some of the key markers to date are presented. For ease of reader access and temporal coherence, this is arranged as two strands: Physiological Inquiry (see sections 2.2-2.4), and Neurological Inquiry (see section 2.5). Naturally these evolved simultaneously and each informed the other, and it is hoped that within this presentation format these can connections can be simply made.

Throughout the 20th Century three major theoretical Psychology Frameworks have developed around the emotions that serve as the grounding for practical inquiry. The most often cited origin of modern emotion research stems from Philosopher and Psychologist William James's 1884 publication 'What is an Emotion' (James, 1884). Against his peer's popular considerations of the emotions being a cerebral process, James proposed that emotion was simply the perception of physiological changes in response to external stimuli. If one were to encounter a dangerous situation, then the bodies' autonomic responses would present the necessary internal conditions required to respond to the stimulus. In the example of danger, a subject may flee and then perceive his or her emotional state as fear. Whilst the Jamesian theory spiked much controversy, it laid a formal marker for further hypothesis testing of; what is emotion? , How and why it occurs, and importantly how it may be detected through scientific investigation.

Following James, two major theoretical investigative strands of thought quickly developed. These two contrasting Psychology models were, firstly the division of the emotions into discrete states such as; happy, sad, or angry, and secondly the envisaging of emotion as a spontaneous continuum. In the latter parts of the century, spearheaded by the introduction of digital technology and its possibilities, a third framework emerged citing a more holistic view of the human organism. This was an appraisal based componential model.

2.2 Theoretical Frameworks: Discrete

Discrete theories of emotion can be seen to align with the deterministic ideas that were prevalent at the time of its formation. Within Determinism, all phenomena are considered to be reducible to fundamental discrete units with mechanistic qualities. In the same way, discrete theory views the emotions as a form of architecture built upon a fundamental core set of emotions. These are widely termed the 'Basic Emotions' or the 'Peak States', and their varied compositions may lead to instances of further complex emotions arising.

The view is upheld that these basic emotions are determined by underlying physiological signatures and by detecting their variations one may be able to gain an objective insight into their nature and occurrence. A range of sensors have been proposed and adopted in this approach measuring a wide range of physiological signals, for example: pulse, skin conductance, facial expression, blood pressure and blood sugar levels, and pupil dilation. Of these, the most prevalent is emotional facial expression recognition, thus this text will centre on its development to highlight this theory. It is important to state that the detection of the emotions through the above-mentioned sensors only provides access to the second order of emotion, its expression, and not of its generating point or function.

Writing contemporaneously to William James, and as an extension of his 'Evolution Project', Charles Darwin (1934) published accounts of his natural observations of species wide similarity between external gestural expression and its association to internal states and feelings.

Darwin developed a protocol of testing these ideas within human populations. A series of photographs featuring a number of facial expressions were shown to individuals. The repeated successful recognition of these solidified his belief that facial expressions were biologically determined in correspondence to internal states. These fitted in with his wider theories of evolution and natural selection.

Almost half a century later Flloyd Allport echoed these ideas of biological hardwiring. In support of the Jamesian model, he suggested a foundational set of discrete emotional states that were triggered by current physiological configurations. In the 1920’s Allport published his accounts of a mechanical one-to-one correspondence between facial muscles (expressions), and emotional states. Allport proposed a bipolar Valence structure to the Autonomic Nervous System's (ANS) activities, which was only able to distinguish between pleasant and unpleasant conditions. Here the Sympathetic Nervous System (SNS) was triggered by negative states, and in turn, the Parasympathetic Nervous System (PNS) for positive states (Grendon & Barrett, 2009). He emphasised that the numerous facial expressions may be distilled to a small set of six fundamental emotional states which he saw as; pain-grief, surprise-fear, anger, disgust, pleasure and neutrality. Further he pursued the argument for the attachment of keyword labels to the emotions for simpler distinguishability.

These sentiments of automatically triggered physiological protocols by hardwired emotions in response to external stimulus were furthered in Psychologist Silvian S Tomkins 'Affect Theory'. For Tomkins (1982), the emotions are the most basic motivational forces of human life. He believed they comprised of a binary nature, which he arranged in a polarity pair list: of positive (interest or excitement, enjoyment or joy, startle or surprised) and negative (distress or anguish, fear or terror, shame and humiliation, contempt, anger or rage).

In a reversal of the Jamesian proposal, Carol Izard (1991) veered towards a cerebral centre of emotion generation, where each emotion was housed within a particular neural network, producing the relevant corresponding response. It was during Izard's first-hand subjective experience of parenting where he evidenced a base set of emotions in a growing newborn, before full cognitive competence. From his derived insights Izard sought to develop an framework of 10 discrete basic emotion categories comprising of; Interest, Joy, surprise, distress, contempt, fear, shame, disgust and guilt.

Figure 2.1: Illustration of the six discrete emotions as proposed by Paul Ekman & Wallace Friesen. Clockwise from top left; anger, fear, disgust, sadness, happiness, and surprise. (Ekman et.al, 1975).

Working alongside Tomkins, Paul Ekman and Wallace Friesen continued to develop their foundational framework for facial emotional expression recognition. This evolved to become a comprehensive and prominent form of emotional expression detection. Through highlighting the underlying arrangements and possible configurations of the 42 facial muscles, they were able to develop a practical taxonomy of facial emotional expression. Echoing previous sentiments in similar inquiries, Ekman strengthened the proposition that real felt emotions primarily reveal themselves through facial expression, and further that body posture revealed how a person may be coping with that emotion (Ekman & Friesen, 1975).

Ekman's studies across technological and remote cultures upheld his theories leading to the general acceptance of their universality. In a method similar to Darwin's, Ekman presented photographs of his six prototypical emotions; happiness, sadness, surprise, fear, disgust and anger, to gauge participants recognition of these states (see Fig 2.1). This successful method received peer wide acceptance, and was developed into a system entitled the 'Facial Action Coding System (FACS). This is the most widely recognised and utilised body of research in the field of emotion recognition.

However alongside endorsements, there were a number of criticisms of his theory and method. His prototypical emotions were not considered functional in the everyday, and were seen as exceptional states rather than the norm. Further that they failed to account for the rich variety of emotional experiences. Finally, the ecological validity, experimental design, and results were not safe from question and criticism. (Russell, 1994, Russell 2009)

Ekman's response to these claims of wider emotional experiences, led to the extension of his theoretical envelope through introducing the notion of 'Emotional Families'. Here emotional blends operated as transitional compensations for emotional diversification.

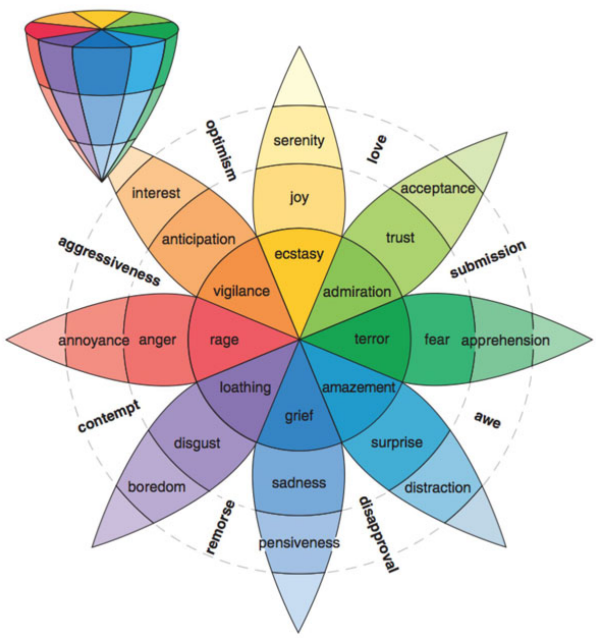


Figure 2.2: Robert Plutchik's emotion wheel depicts how complex and subtle emotions may emerge from 8 basic emotions [7].

There are similarities in this to Psycho-Evolutionist Robert Plutchik's model. Plutchik (2001) postulated how a primary set of eight bi-polar discrete emotions; joy-sadness, anger-fear, surprise-anticipation and trust-disgust, functioned in a form of continuous space to compensate for continuous state transitions and secondary

[7] Retrieved from http://www.6seconds.org/wp-content/uploads/2011/06/Plutchik.gif

emotional blends. Reminiscent of Newton's and Goethe's colour wheels, Plutchik created an metaphoric visual representation through a unfolded two dimension colour wheel, which could be reassembled into a three dimensional cone (see Fig 2.2). He used colour as a metaphor for the state arrangements, utilising tone to represent intensity levels, cross sectional complementary colour for emotion similarity, and contrasting colour for bipolar states. Here different locations within the space presented the full emotional palette. For example, a location that is a compositional blend of disgust and anger, equates to loathing, rage, hatred, and then hostility.

This archetypal model of a small set of discrete emotions has been applied in many experiments with a range of sensors, with each experiment having to develop their own methodological protocols.

2.3 Theoretical Frameworks: Dimensional

The second Psychology Framework to emerge in emotion research opposes notions of a basic set of hardwired emotions. Rather it considered the emotions to operate within a spontaneous and continuous space. Here, Peak States and a wider set of emotions may be metaphorically incorporated as Islands relating to one another in a systematic fashion within an all-encompassing space. Within this theory the emotions are constructed through underlying dimensional features such as Valence, Arousal, and Dominance. Each dimensional vector can be considered to have iterative steps from a negative index to a positive index. The composite of such dimensions are believed to contribute to a felt emotion, its expression, and its detection. This model takes into consideration the more subtle ranges, spontaneous occurrences, and the temporal elements of emotion in natural experiences.

In the 1970's Psychologists Albert Mehabrian and James Russell (1974), proposed the PAD framework for the continuous measurement of emotional states. The acronym PAD signifies its tertiary scalar dimensions of; Pleasure – Arousal – Dominance. Each dimension at its outer reaches represents either a positive or negative value. The Pleasure dimension measures Pleasing to Unpleasant, Arousal measures intensity of Arousal through to Non Arousal, and the Dominance vector ranges from feelings of Powerfulness to Submissiveness in regards to the stimuli, which is inclusive of environment. It was felt that these factors were the components that mark up the emotional experience.

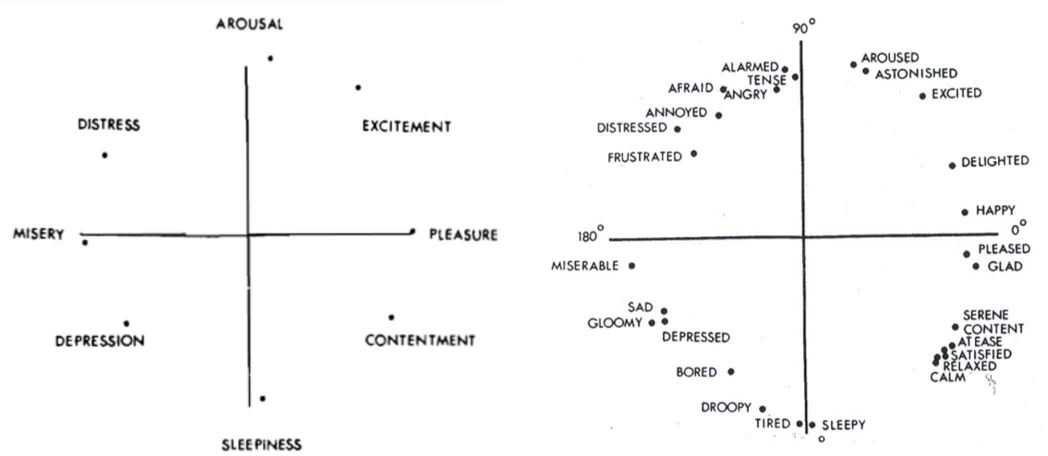


Figure 2.3 Russell's 2-dimensional circumplex models; constructed from Valence and Arousal vectors [8].

In the following years, Russell (1976) drawing on the work of Scholsberg (1954) further collapsed this model to an easily presentable two-dimensional circumplex model. Here, Russell plotted only Valence and Arousal to allow their combination to correlate and flexibly pass through a range of Peak States. Around the circumference of a circle James plotted the emotional conditions correlative of the variables of these two components (see Fig 2.3). Russell’s circumplex model is widely used in emotion research, and both of these dimensional models demonstrate a space where additional confirmed vectors may be added, and further proposed states mapped to a location.

The strong opposition and criticism of its descriptive lack of any subjective appraisal in emotion was somewhat rebuked through the expansion of this model. Lisa Fieldman Barrett extends the dimensional model to account for wider neural and physiological processes in what she terms the 'conceptual act model' (Barrett, 2009, Lindquist et. al. 2012). This Psychological Constructionist approach draws upon the assumption that the emotions (or situated conceptualisations) are psychological events that emerge from more principle psychological operations, which are not unique or specific to emotion, but share a commonality with processes not involved with emotion.

The model comprises of three major processes. (i) ' Core Affect' is considered as the mental representation of bodily sensations associated with the vectors of Valence and Arousal that operate with motivational functions for the organism such as, approach

[8] Retrieved from http://www.wwu.edu/culture/images/altarriba.jpg

or avoid. (ii) 'Conceptualisation', is the comparative processing between previous experiences, stored emotional memories and incoming perceptions, from which the meaning of a immediate sensation or event may be constructed. This process of making meaning out of core affect is assisted by (iii) 'Executive Attention' which functions as a form of filter. This control device for the whole process determines the utilization and suppression of a plethora of representations that may be made available in the moment of making sense of perceptions.

Similarly For Russell, Psychological Construction is the progression of the dimensional model to counter-act criticisms of the exclusions of subjective decision-making. Core affect is the response of the two dimensions of Valence and Arousal as a process intertwined within perception as a form of decision-making that is not necessarily conscious, it is an underlying neurophysiological state influenced by these two vectors. Core Affect operates under the umbrella processes of Psychological Construction whose collective processes may lead to emotion regulation, subjective emotional experiences, the expression of emotion, and also the association between these aspects. (Russell, 2009)

2.4 Theoretical Frameworks: Appraisal

A third theoretical Psychology perspective was to emerge in the latter part of the 20^{th} century. It highlighted the centrality of cognitive processing in a networked elicitation of emotion. This approach was structured around the idea of cerebral appraisal. This novel approach is both informed and driven towards the mapping of human experience in the construction of artificial machine intelligence through a computational approach. In this way further understanding of the emotions may be gained by building synthetic emotional representations from the ground up, iteratively driven forwards by sensing aspects that are lacking in current synthetic system.

Magda Arnold is cited as one the initial proponents of this theory (Plutchik, 2003). She was amongst the first to formally consider emotional responses as arising from a series of cognitive evaluations or appraisals to stimuli. Whilst Arnold's publications were widely derided for their incompletion and she experienced strong undermining due to her gender, this new approach served as a foundational platform for subsequent generations to build upon. In the 1980s due to the sophistication, miniaturisation and proliferation of computing technology appraisal theory began to become prominent.

Psychologist Richard Lazarus (1991) saw the emotions residing in discrete categories, which were able to contribute to further states. He cited emotional states as arising from a structured appraisal process of ‘cognitive-motivational-adaptive-and physiological activity’, in an individual in relation to the environment. Here, the appraisals were subconscious or autonomic to the subject’s survival; i.e. either harmful or beneficial. In a hierarchical system he saw primary appraisals concerned with the relevance of interaction to the achievement of personal goals i.e. ‘thwart – achieve’, and secondary appraisal in the accountability of ‘blame or credit’, which heavily affected ones coping potential and of future expectation. Further Lazarus highlighted coping mechanisms within his theorem. This consisted of direct action and subsequent emotional coping. In summary, Lazarus proposed the emotions as a form of responsive script-narrative.

Nico H Frijda aligned to both Arnold's and Lazarus's propositions. His term 'action readiness’ condensed the emotions to goal oriented motivations where achieving a goal brought a positive emotion, whilst negative emotion arose in instances of a threat to the organism or its goals. His text the 'Laws of Emotion' (Frijda, 1988) emphasises this system through hard-wired biological deterministic physiology, which is triggered through personal appraisal. This appraisal is dependent on the personal history of the organism and as such accounts for why different basic emotions and different secondary mixes of emotion may be elicited in different subjects to the same phenomena, as it is always in relation to personal relevance and meaning. Similar to the Jamesian consideration they prepare us for action, but through appraisal. His term 'action tendency' succinctly states this; the emotions are the tendency for a subject to engage in behaviours central to their own concerns and needs.

Klaus Scherer (1984(a), 1984(b), 2009) further advanced this field through his appraisal based componential framework for both recognising and expressing emotions. Its consolidating multi-modal approach attempted to account for issues of multiple subjects’ variant emotional responses to the same stimulus. Scherer’s appraisals consist of several components which are cited as; (i) cognitive appraisal or evaluation of stimuli and situations, (ii) physiological activation or arousal (iii) motor expression, (iv) motivational intentions, (v) subjective state feeling.

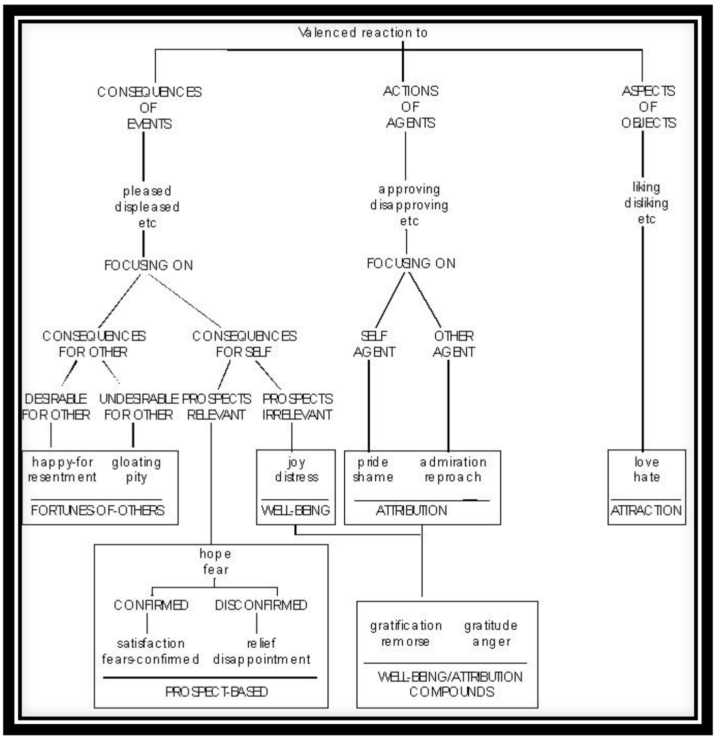


Figure 2.4: The Ortony, Collins, & Clore's appraisal model caters for Valence model focuses on tertiary branches of engagement with; agents, events, and objects [9].

In a similar approach Ortony, Collins, and Clore (1988) present a framework where recognition and expression could be united in a reversible system (see Fig 2.4). There basic assumption moved from focusing on emotion per se, to the perceived engagement with the world. There Valence based model focused on decisions to tertiary branches of engagement with; agents, events, and objects.

"We define emotional intelligence as the subset of social intelligence that involves the ability to monitor one's own and others' feelings and emotions, to discriminate among them and to use this information to guide one's thinking and actions."
Salovey & Mayer Pp 189.

Figure 2.5: The Salovey & Mayer definition of Emotional Intelligence.

As mentioned above many aspects of the appraisal model may be seen as particular to the fields of synthetic and constructible intelligences. The regard of incorporating emotion as a central element of machine intelligence followed Salovey

[9] Retrieved from http://www.ruebenstrunk.de/emeocomp/36fa2f7a.jpg

and Mayer's (1990) paper 'Emotional Intelligence'. They updated themes of social intelligence as a requisite for intelligent computing (see Fig 2.6). In there thesis they outlined four branches of this; perceiving, reasoning, understanding, and managing.

Daniel Golemans (1996) expansion and popularisation of this borrowed term 'Emotional Intelligence ' re-contextualised the importance of the emotions as a form of intelligence within social circles. The popularity of his writings led to these ideas being incorporated as an essential skill-sets for echelons of the corporate world. Both Salovey & Mayer, and Goleman demonstrate the need to consider emotion as an inherent part of any intelligent communication.

One of the seminal texts in the consideration of emotion as central to intelligence and communication is Rosalind Picard's thesis 'Affective Computing' (1997). This generated a new term for this field of study, and echo's the sentiment of Salovey's proposal. Picard's vision pre-empts computing that relates to, arises from, or deliberately influences emotions, presenting a form of computing with abilities to recognise and express emotions. This is with the aim of more rewarding and engaging human-computer-interaction experiences. The scope of Picard's research is far reaching. In what may be called her manifesto, she considers all aspects of Affective Computing from its inception through to predictive forms of technology and scenarios. Researchers and research groups have taken many of these batons on board. One such example is the Semaine Project (McKeown et. al, 2012) where artificial emotion coloured characters engage in a form of simulated dialogue which is responsive and provocative to the user. The system uses facial expression recognition and voice intonation algorithms in a form of simulated exchange.

2.5 Theoretical Frameworks: Cognitive

The James-Lange (James, 1884) physiological model of emotional is the most often cited beginning in all areas of modern emotion research. Here it was envisaged that emotion functioned as the cognitive awareness of sensations accompanying physiological changes in response to perceived stimuli.

Flaws in this outlook were first highlighted by Cannon & Bard (1927,1928), who questioned anomalies of the commonality of particular physiological responses between variant emotional states. Through a series of feline ablation experiments, they demonstrated that emotional expression could still be perceived despite regional cortical damage and removal, as long the hypothalamus remained intact. This summary contradicted James's view of the dependency of the operating perceptive brain (cortex) to perceive physiological changes. They put forward their main proposal, whereby bodily changes occurred simultaneously to cognitive processes in emotional experience and expression. They cited the thalamic region as playing a major role in the experience and expression of emotion.

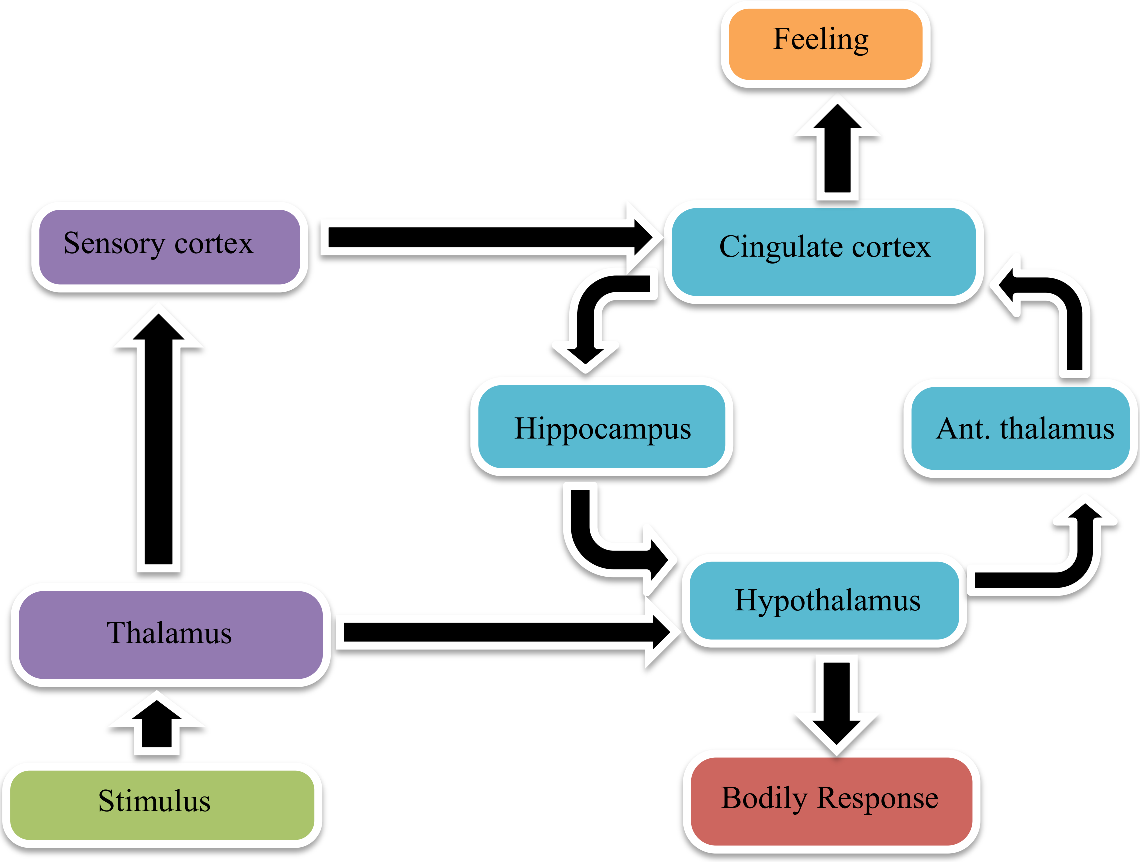


Figure 2.6: Illustration of James Papez Neuro-anatonomical Circuit.

James Papez (1937) extended the focus on the thalamic region through his Neuro-anatomical circuit (see Fig 2.6). Here sensory perceptions were routed via the thalamus along two main streams. One travelled upstream (somatic) to the cortex for rationalising feelings and memory encoding, whilst the other was diverted downstream (visceral) producing physiological responses.

This circuitry included the idea that 'Thought' could drive the circuitry in a top down movement, to invoke and regulate physiological feelings and emotions. Papez cited the Hypothalamus, Cingulate Cortex and Anterior Thalamus, as emotional circuitry through his experiments of injecting the rabies virus into feline brains and observing their progressive destruction.

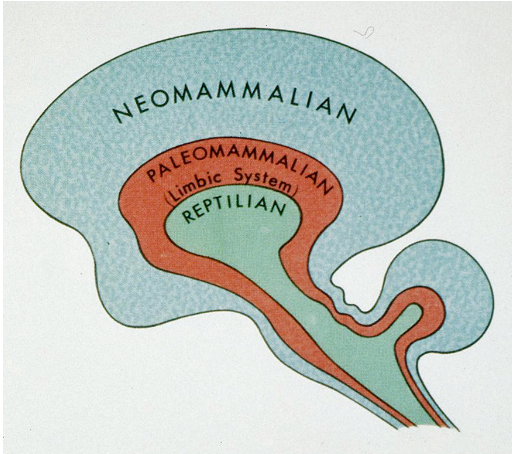


Figure 2.7: An illustration of Paul Maclean Triune brain. His model is structured from three interconnected and inter-communicative evolutionary steps [10].

In the middle of the 20th century Neuroscientist Paul Maclean attempted to clarify why sub-cortical regions play such a major role in species-similarity emotional behaviours through his Triune brain model (Maclean, 1990). For Maclean the human brain is constructed from three interconnected and intercommunicative evolutionary steps, each having its own functional processes and memory capabilities (see Fig 2.7).

According to his model, the earliest section of the brain to evolve, which includes the brainstem and the cerebellum, is termed the Reptilian Brain. It is seen to contribute compulsive behaviours, ritualistic responses and controls autonomic functioning of the inner organs. On top of this, is the intermediate or Palelomammalain Brain, which Maclean entitled the Limbic System. This included the hippocampus, which he saw as the centre for emotion. The limbic system is shared species wide in mammals and is concerned with instinctual motivational drives such as emotion, reproduction, feeding, and self-preservation.

The final most recent evolutionary step of the brain is the development of the neo-cortex or neo-mammalian brain. It is found in most animals but developed to a high degree of sophistication in primates and especially humans. It is made up of two

[10] Retrieved from http://mybrainnotes.com/triune-brain-theory.jpg

hemispheres, which are wired for varied processing. Essentially, it allows for higher modes of mental activity, such as logic, abstraction, reflection, rationalisation and complex communications such as language. Maclean proposed it was the comparative signals of physiological changes alongside the context of the perceived world, which generated emotional experience.

In their 1960's theoretical psychology model, Stanley Schacther and Jerome Singer (1962) aligned to Maclean's view in their presentation of a dual-factor approach with emphasis on cognitions role. They outlined the importance of context to emotion, whereby a person experiencing unattended physiological cues would search for the causality of their generation. This may align to previous cognitions, interpretations of experience, or even the surrounding environment. Their behavioural experiments comprised of groups of participants being administered with either a stimulant or a placebo. Whilst seated in a waiting room, a 'stooge' would perform a series of provocative actions as they filled in a questionnaire. This asked for increasingly absurd, personal, and emotionally provocative information. Schacther & Singer analysed these questionnaires assessing the impact of the 'stooge' on their levels of emotion and the outcome of subsequent moods in their responses. Thus, Schacther and Singer suggested visceral responses were a primary factor, whilst the cerebral contextualisation provided the secondary factor.

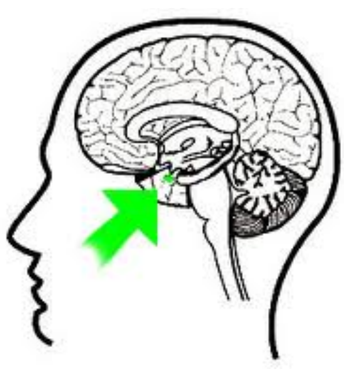


Figure 2.8: Joseph Ledoux's experiments demonstrated the Amygdala's involvement in the emotion of fear, showing how different regions of the brain may contribute to different emotions [11].

Whilst many of the regions outlined in both Papez's circuit and Maclean's Limbic system have been confirmed, they provide only a partial and broad description

[11] Retrieved from http://alt-sites.tripod.com/pictures/level05amygdala.jpg

of the brains role in emotion. A new detailed understanding of how the different regions of the brain correlate to different emotions first emerged in the work of Joseph Ledoux (1996). Ledoux combined Pavlovian fear conditioning experiments with rat ablations, to isolate the amygdala as the region concerned with the processing of fear and fear responses (see Fig 2.8).

Reminiscent of Papez's streams Ledoux discovered that sensory information was directed along two roads, which he called the high and low roads. The low road sent crude archetypal information to the amygdala, which instigated a series of relevant physiological responses to circumstances of survivalist threat. Approximately 12 milliseconds later, detailed information would arrive at the cortex for a form appraisal with the potential to either dampen the response to the perceived threat, or to respond accordingly. The hardwiring of this is seen as a biological survivalist mechanism, as sustained procrastination solely via the cortex may render the organism obsolete. What this further revealed was that memory is distributed, and functions differently across the brain.

'Affective Neuroscientist' Jaak Panksepp (1988, 2000, 2010) explores the foundational neural nature of emotion across species with an interest in exploring how the emotions are generated. As an advocate of the evolutionary brain theory, he also cites emotion as residing within its most ancient subcortical regions. His laboratory experiments focus on motivational processes, social emotions, and behaviours in non-human subjects, which he believes apply across species. Thus, Panksepp cites a very different set of base emotions/affective states that all have motivational forms of biological programming from which further emotions evolve in relation to the complexity of their found environment. He cites seven primary fundamental social emotion-affective states, which are; seeking, rage, fear, panic/loss, play, mating, care. He views these as essential communicators within communal species. Of particular interest are 'seeking' and 'play', which challenge the notion of solely reward/goal based motivation theories and approaches to emotion. Panksepp's work stands out in emotion research as it attempts to go beyond the neuroanatomical mapping of emotion, attempting to unveil its underlying function.

Modern neuro-imaging techniques such as Magnetic Resonance Imaging (MRI), functional Magnetic Resonance Imaging (fMRI), Computed Tomography (CT) , Positron Emission Tomography (PET) , Electroencephalography (EEG) and Magnetoenchephalography (MEG) have allowed for an unprecedented level of access to brain function and neuroanatomical observation. They have assisted in the progression of the cognitive framework whilst also side stepping ethical issues of using laboratory mammals and primates.

In 2000 Antonio Damasio et. al (2000) used a PET scanner to observe the neural networks triggered by different emotions. The enlisted participants were asked to recall experiences relating to sadness, happiness, anger or fear, following being injected with a form of radioactive dye that attaches itself to glucose in the brain. As brain activity commences it draws in the glucose with the attached dye for fuel, which can be traced through the PET scanner. Damasio and colleagues were able to visually reveal that different regions of the human brain are involved for variant emotional and affective states (2000).

Damasio's (1990) early experiments with cortex damaged patients led to his somatic marker hypothesis. This is a proposal for how emotion biases behaviour and rational decision-making. Damasio hypothesised that physiological reactions are tagged with emotional significance and encoded within memory forms, which then become available as influential Valence inclusions in future perceptions and rational appraisals.

He engaged groups of participants of whom some had damage to their cortex's and others not, in the 'Iowa' gambling task where-by low risk turn taking presents higher rewards than high risk turns over the same number of turns. Damasio was able to demonstrate how cortex damaged participants were unable to learn form previous turns and feelings. In contrast participants without damage were able to follow their gut instincts and eventually understand that selecting the low risk option would prove more fruitful. This highlights both emotion recorded within memory envelopes and how gut feelings or emotion play semblance in our decisions which are there not purely rational, but influenced by our own Valances. Damasio asserts in the same manner that Ledoux does, that pure rationalisation as an appraisal function would impinge on arriving at a conclusion.

Richard Davidson is one of many pioneers in the neuroscience of emotion research. Since the 1980's Davidson's research focus has been directed towards uncovering the relationship between emotion and cerebral activity. Having conducted some of the earliest experiments into the detection of emotion through EEG (see chapter 3, and for a fuller exploration of EEG usage in emotion detection.) Davidson's focus has expanded to consider some of the wider facets of the emotional experience, inclusive of accounting for subjective variant responses to specific stimuli. This considers the structure from which emotion and mood arise. Whilst this has been called temperament and personality trait, Davidson’s has named his model ‘Emotional Style’ (Davidson, 2012).

The model is constructed from six variable axis representing; Resilience, Outlook, Social Intuition, Self-awareness, Context, and Attention. Each dimension accounts for a construct of an emotional foundation from which a propensity of emotional elicitation can be predicted. Further Davidson has demonstrated that this foundation is permeable. In the same manner that the brain can rewire itself through what is known as neuroplasticity an individuals emotional language can be altered through directed learning. This shows that emotion may not hardwired by nature in our genetics, but rather as Davidson appeals that nurture is nature!

2.6 Conclusion.

In the above passages of text the progression of the field of emotion research from its humble origins have been laconically laid out. Through formal scientific inquiry a number of theoretical windows and access points to the emotions have emerged. Whilst these frameworks have attempted to map out a overarching consensual theory of emotion; unveiling its underlying facets, origins, locations, processes and functions, these attempts have been combated by the complexity of the subject matter. This complexity of interlinked near-simultaneous multiple processes, dispersed throughout a holistic organism in both linear and non-linear feedback fashions, of which some may be regulatory, serve only to make these investigations by reduction seemingly more problematic.

This is confounded by the inclusion of hidden and private neural spaces. The emotions role and involvement in memory, perception, decision making, rationalisation and motivation further assist the difficulty in the precise pinpointing of this central phenomena. Pankseep's suggestion of further forms of emotion yet to be considered can only add to this. Thus the pioneers making advances in the field however partial, and those building upon their foundations should be heralded for their achievements and for engaging with this difficult register of knowledge.

The aims of this particular project are to explore a process whereby we may be able to gain an insight into the subjective experiences of the viewing of cultural art forms via the emotions, and in turn how the emotions may be used as an artists material for makers. When we take into consideration an individual viewing an artwork, we may state that whilst an initial immediate response may be apparent, this is a part of a temporal activity that may be framed within a solitary experience. Thus this may be of an unrevealed or less shared nature where emphasis is on internal feelings and cerebral experience and not necessarily reflective of outwardly expression in the same way that prototypical or social emotions may occur. In consideration of this, the most beneficial model to meet our aims is the dimensional model. Whilst this may only present limited information, through it we may be able to detect whether a viewer may be experiencing positive or negative response sensations, and also the entwined level of Arousal; relaxed or excited. From this we may be able to infer transitional emotional states rather than a particular prototypical emotion. However it may also be of use to incorporate the discrete model within our annotation methods, and this can be tested in experiments to compare the models suitability.

Neurologically, we have already defined that EEG is the most desired and viable route, both in terms of access, finance and mobility. Richard Davidson's oeuvre shows a potential framework to match the context of this projects aims, and thus in the next chapter EEG and its relationship to the emotions will be explored.

We began this chapter with a passage from Daniel Goleman citing the emotions centrality to our experiences of being in the world. At the outset this may have seemed slightly outlandish, yet having mapped the terrain of emotion research his passage morphs into a reflection and condensation of all that has been discovered and articulated in the past 100 years.

This literature review was an essential aspect in this research as it assisted in the grounding of the project giving directions to its trajectory. It contextualised its outer reaches, its limitations, and also its potentials modes of inquiry.

CHAPTER 3

Electroencephalography & Emotion Detection

3.1 Introduction.

Prior to the emergence of modern neural imaging technologies, insights into neural circuitry and function were obtained solely through Psychology and behavioural studies. These heavily relied upon random instances of brain injury and adverse neurological dispositions for opportunities of insight.

The most often cited example is of railroad foreman Phineas Gage. In a accidental on-construction-site explosion, his left frontal cortex was impaled by an iron rod. Although Gage survived the accident, his behavioural and emotional disposition morphed towards increasingly negative and aggressive tendencies. Posthumous studies on Gage showed mass damage and obliteration of this region leading to the general theory that linked the frontal cortex to emotional responses and engagement. Subsequent studies continue to consistently support this.

The first scientific apparatus to provide methodological access to the brains inner workings was the Electroencephalogram. Its origins can be intertwined with the historical timeline of electromagnetism that stretches back to 1600 AD. In the 18th Century whilst experimenting with deceased amphibians, Luigi Galvani noted their consistent muscular contractions when exposed to an electrical current. This accidental discovery of the electrical basis of nerve impulses, led to supposition of the 'Electrical Brain'.

Many electrical experiments with mammalian cortex's continued throughout the late 19th Century, namely by Vasili Danilevsky in 1876, Fleischel von Marxow in1883, Adolph Beck in 1890 and in the 20th Century; Vladimir Pravdich-Neminsky 1913 (Swartz and Goldensohn, 1998) and Richard Canton in1924, (Ormerod, 2006)

3.2 Invention: Hans Berger.

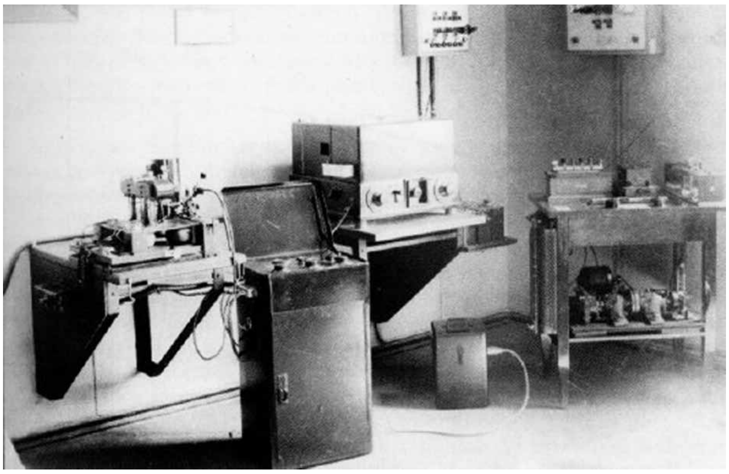


Figure 3.1: An image of Hans Berger's Laboratory (1926-1931) at the University of Jena where Electroencephalography was conceived [12]. Berger used the double coil method normally used in electrocardiography, as the basis for EEG.

Psychiatrist Hans Berger progressed his own 'Electrical Brain' experimentation with mammalians to begin incorporating human subjects. Berger sought to develop a practical neural equivalent of electrocardiography. In 1929 he succeeded to publish a paper that introduced the official term for this process, 'The Electroencephalograph' (EEG), which he defined as a technique for detecting and recording the rhythmical electrical wave oscillations generated by the brain (translated by Gloor, 1969)

Berger's work was driven by a personal and philosophical interest in unlocking "the secret nature of mans nature as a psychophysical being", a belief that strongly contested the deterministic view of human duality. In over 101 recording sessions with 38 different patients suffering from either trephine openings or skull deformations, Berger was iteratively able to formulate a non-invasive method whereby the brains electrical activity was rendered visible. Berger successfully used lead foil electrodes sandwiched between flannels that were held firmly in place on the scalp with rubber bandages to obtain the signal. This successful technique formed the basis of modern EEG and some of his first documented recordings were conducted on his 16-year-old son Klaus.

[12] Retrieved from http://www.mpiwgberlin.mpg.de/resrep00_01/images/Jahresbericht_img.large/135.jpg

Through his experimentation, Berger consistently noted two spontaneous and regular continuous rhythmical currents oscillating in the frequencies regions of 10-11 Hz and 20-30 Hz, which he labelled 'Alpha' and 'Beta'. Through rigorous elimination, he was able to discern that these signal were not arising from physiological artefacts. Further experiments led to the finding that these frequencies became animated in response to neural activity and that such activity led to a decrease in alpha. These demonstrated how mental engagement and could be used in an inverse relationship as a measure for cerebral activation.

At the time, Berger's work was largely overlooked, dismissed, halted, and his laboratory was disassembled due to both his political beliefs and his peer's perceptions of his training as Psychiatrist rather than a Neurologist. Yet Berger was able to form the basis of major non-invasive EEG clinical techniques for detecting epilepsy and neurological disorders.

3.3 The Language of the EEG: What's was Berger detecting?

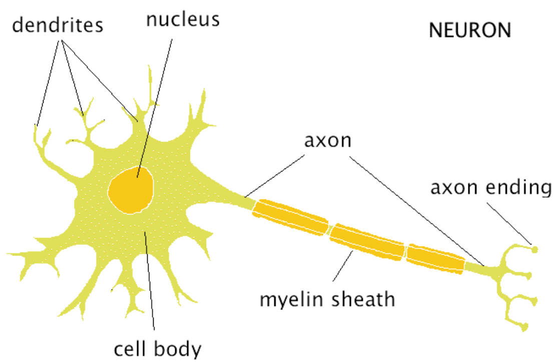


Figure 3.2: An illustration of a single brain neuron.

The human brain is constructed from billions of minute cells known as neurons. Each neuron consists of a nucleus surrounded by a cell body from which extend several branches called dendrites, and a singular long branch called the axon. Each dendrite contains embedded receptors with which it can receive signals from other cells via their axon, which has transmitters at it terminus (see Fig 3.2).

The signals (nerve impulses) are delivered in the form of chemical particles known as neurotransmitters. These travel in a wave formation called an Ion which carriers an electrical charge. The exchange of chemical particles occurs across a minuscule gap known as the synaptic cleft, and with sufficient charge is passed along to further connected neurons. Although a single charge is too small to be detected, when millions of neighbouring neurons fire together they generate a waveform that can be detected by an EEG electrode. Whilst the amplitude of the signal is of a high value on the cortex, on the scalp it is less pronounced (in millivolts), hence forms of signal amplification are a necessity.

This issue of amplitude taps into both the major advantages and disadvantages of EEG. Its main advantages are in its non-invasive nature and its high temporal resolution that allows brain activity to be accessed in timeframes upwards of 20,000 Hz (Carter, 2009). Its major disadvantage is its low spatial resolution. The presence of the scalp between electrode and cortex creates a blurring of the signal. This low signal to noise ratio eliminates the potential of any detailed anatomical study within current EEG technological systems.

3.4 Contemporary Standardized Measures: 10-20 International System.

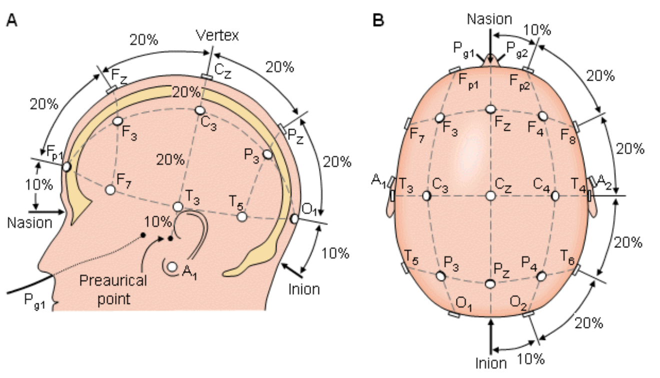


Figure 3.3: The 10-20 International system for EEG electrode placement [13].

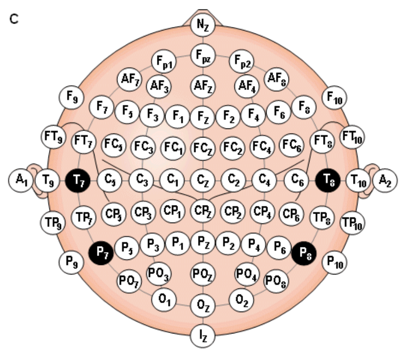


Figure 3.4: Further electrodes can be added at 50% ratios [14].

The International Federation of Societies for Electroencephalography and Clinical Neurophysiology has endorsed a universally recognized system for electrode placement,

[13] Retrieved from http://neurologiclabs.com/wp-content/uploads/2013/12/EEG.gif
[14] Retrieved from http://neurologiclabs.com/wp-content/uploads/2013/12/EEG2.gif

entitled the 10-20 International system, which fragments the scalp into particular sections.

The distance between the Nasion and Inion points is divided into percentage blocks of 10 and 20 percent. Measuring 10 % from both the Nasion and Inion points is where the electrode placements begin and the subsequent electrodes are placed at 20% distances (see Fig 3.3). Whilst initially the 10-20 system compensated for the arrangement of 21 electrodes, additional electrodes are configurable at 10% intervals between the original electrode positions (see Fig 3.4). Further electrodes are often added for additional ocular and muscular movement for both multi modal detection, and signal-noise elimination (Sansei & Chambers, 2007).

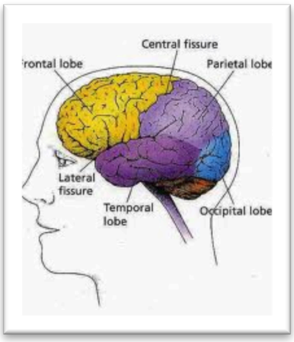


Figure 3.5: The Different label indices in the 10-20 International system relate to their proximity to each of the brains lobes.

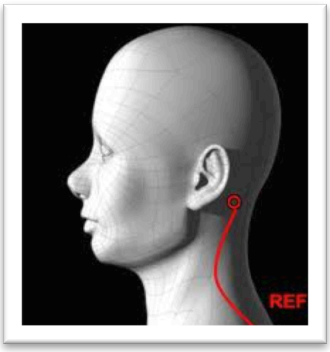


Figure 3.6: The mastoids, (behind the ear) are commonly used as a reference channel due to the low electrical activity of this site [15].

The Label Index for an electrode position relates to the cerebral lobe of their placement (see Fig 3.5). Hence, F =Frontal Lobe, P = Parietal Lobe, T= Temporal Lobe, O = Occipital Lobe, C = Central, FP = Prefrontal and A, Ear (ground/reference). Numerically odd numbers are for the left hemispheres and even numbers for the right hemisphere.

As each EEG electrode yields numerical differences between the electrical activities from two electrode locations, there is a need for a common reference channel that can be used for all the configured electrodes (Hagemann, Naumann & Thayer,

[15] Retrieved from http://stimlab.org/wp-content/uploads/2012/05/meg-ecg-refleads.jpg

2001). The most commonly used reference positions are; (i) CZ (see Fig 3.4) which is central to both hemispheres and also nansion and inion points, (ii) a double configuration placed either on both ear lobes or linked mastoids (see Fig 3.6). The ears and mastoids are considered to be locations where there is the least amount of electrical activity, thus offer a clearer reference channel. The double configuration further provides a conceptual centralised located reference point, which further eliminates hemispheric bias. (iii) Average Referencing (Ar) entails dividing the sum of all channels (excluding the electrode concerned) by their quantity. This method would be used for each electrode. Ar, can also be used in a configuration called montage where a body site such as the wrist or leg is used, in this situation all the head electrodes are averaged as mentioned above.

Two major trends have emerged in the neural detection of emotion through EEG. As with the psychological models of emotion we can see two major temporal factions analogous to psychological discrete and dimensional models. Discrete measures often relate to Event- Related Potentials (ERP) which mainly follow a technique of averaging multiple epochs against the same response for a minute timeframe. Hence ERP can be used to gauge responses to a single event type stimulus. The secondary method of Asymmetric Hemispheric Difference (AHD) can be seen to be more represented within a dimensional model as it can serve continues readings. As this project it is concerned with EEG in naturalistic settings, and hence continuous timeframes, the subsequent text will focus only on AHD.

3.5 Asymmetric Hemispheric Difference.

Figure 3.7: An illustration of the two hemispheres of the Human brain. The right side relates to the left side of the body and the left side relates the right side of the body.

Asymmetric Hemispheric Difference (AHD) is the Electroencephalographic method of evaluating differences in 'Activity' and 'Activation' in response to a stimulus between the left and right hemispheres of the brain (see Fig 3.7). As Coan and Allen (2004) distinguish; 'activity' is a measure gauge in hemispheric specialization studies, whilst 'activation' within hemispheres is utilized in the detection of affective or emotional perception responses.

Hemispheric specialization refers to the propensities and preparedness of each hemisphere to engage in the particular processing of information or task; for example speech, which is detected in posterior (rear) regions.

Alternatively, affect or emotional responses are shown to register within anterior (frontal) locations (Davidson, 1998). In a survey study by Coan & Allen (2003), correlative tabling of the first 26 published accounts of EEG studies of emotional detection between 1982 and 2000 showed that electrode locations at positions F3/F4 (anterior location) were always used in reported successful detection.

It should be noted that the above two categorizations operate orthogonally, such that parietal or anterior activation do not bias the other, nor the whole hemisphere. Of further note, 'specialization studies' must take into account the handedness of individuals, as right-handers process speech in their left hemisphere, whilst the opposite is true for left-handers. Hence all EEG studies include in their experimental set-up, awareness of participant handedness.

Electrical activity detected by EEG technology arrives in the form of sinusoidal wave patterns from which it is notoriously difficult to extract meaningful information. However through modern signal processing techniques it is possible to translate the data into an interpretable form through the Fast Fourier Transform algorithm. This translates temporal or spatial information into its underlying components in the frequency domain.

Name	Frequency	Associated state
Gamma	*< 30 Hz +*	*High level processing*
Beta	*13-30 Hz*	*Waking consciousness*
Alpha	*8-13 Hz*	*Relaxed*
Theta	*4 – 8 Hz*	*Light meditation*
Delta	*0.5 – 4 Hz*	*Deep sleep*

Table 3.1: The five major brainwave states and associated spectral frequency ranges.

Through this the electrical signals emanating from the brain can be housed into five sectors each relating to their own particular modes and states of being (Teplan, 2002) (see table 3.1).

In the AHD detection of emotion, EEG focuses on the Alpha brain wave. In Berger's original experiments, he found that mental and cognitive engagement led to the decrease in the amounts of detected Alpha. This became known as 'alpha blocking' and this method is still utilized whereby the alpha power in both anterior lobes is evaluated for differences in response to stimuli. Thus, the alpha signal is read inversely, a drop in its value is seen as a signature of the associated hemispheres mental engagement or activation.

A secondary determination for the use of alpha is its distance from the artefact frequencies generated by muscular and ocular movements. Whilst they are still minimally present their impact is less disruptive of the signal than in higher states such beta and gamma and lower states of theta and delta. The algorithm for alpha power hemispheric difference detection is as follows; (Davidson, 1988)

Alpha hemispheric difference = (right Alpha-left Alpha /right Alpha+left Alpha)

Alternatively this can also be calculated on a logarithmic scale, as absolute power density values may sometimes be influenced by individual idiosyncrasies, of skull thickness or brain volume (Sutton & Davidson , 2000) , Thus :

Alpha hemispheric difference = (log [right Alpha] - log ([left Alpha])

The logarithmic calculation presents a uni-dimensional scale where low numbers equate to greater right hemisphere activation (avoidance (negative Valence)) and higher numbers show greater left hemispheric activation (approach (positive Valence)), whilst equal or zero numbers reveal symmetrical activity (Coan & Allen 2004).

A sum of the total alpha power in the frontal lobe using electrodes F3/F4 can be calculated through the following; (Davidson, 1988)

Sum of alpha power = (right Alpha + left Alpha)

3.6 Support for EEG Valence Detection.

In the 1980's Richard Davidson and Nathan Fox conducted a series of seminal studies into the detection of human emotion with electroencephalography. The first in 1982 was one of the first published emotion studies using this method (Davidson & Fox, 1982). It utilized archived film clips of an actress facially expressing happiness and sadness. Here Fox and Davidson attached EEG electrodes to the scalps of ten-month-old babies to observe greater left hemispheric activation in response to smiling faces in two studies [$F(1, 9) = 6.20$, $P = 0.035$] ,($P < .05$, Sceffe test) & studies [$F(1, 13) = 9.09$, $P = 0.01$] ,($P < .01$, Sceffe test). There was no conclusive correlation for sad faces. Their innovative use of infants for their experiment attempt to create a situation where innate human behaviour could be gauged before major learned bias had set in.

In their second experimental set up Fox and Davidson (1986) replaced visual stimuli, for the most primary sense through which a new-born investigates and engages with the world; Taste. For their experiments, they plied 2-3 day old new-borns with sugar water (positive affect), lemon juice (negative affect), and distilled water (neutral-no affect). Fox and Davidson were able to detect left hemispheric activation for positive affect, in comparison to other stimuli.

The following year a further emotionally motivationally charged experiment was conducted which involved 10-month-old babies who experienced a very brief temporary separation (60 seconds) from their primary carer. Within this time frame a non-familiar person approached. Again, correlations between affect and hemispheric activation were confirmed leading to the formulated basis of relating positive affect with left hemispheric activation, and right hemispheric activation with negative affect (Fox & Davidson, 1987).

In further studies, Davidson working in collaborations inclusive of Paul Ekman and Wallace Friesen sought to validate EEG method alongside facial expression recognition methodologies. They found that subtleties of the registrations between real and feigned smiles were neurally detectable (Ekman, Davidson & Freisen, 1990), further through the examination of expressed joy and disgust, they were able to suggest that the hemispheric model was representational of approach and withdrawal tendencies as opposed to discrete Valence levels (Davidson, Ekman, Saron, Senulis & Freisen, 1990). Both Fox (1991) and Davidson (1992) write accounts associating negative affect with withdrawal tendencies and right anterior activation, and the opposite for positive affect, which allows it to be mapped into a dimension model, but not necessarily in a one-to-one association as with the keywords defined by Russell.

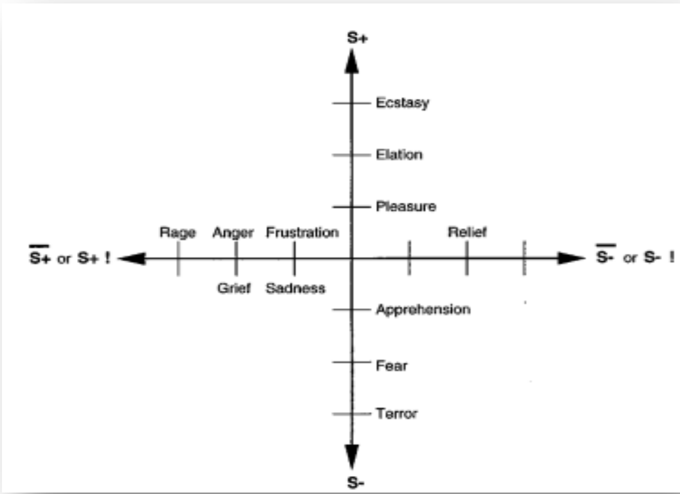


Fig 3.8. An illustration o f Edmund T Rolls approach and avoidance model that allows for the charting of key and subtle emotional states (Rolls, 2001).

Thus, whilst affective Valence within hemispheres can be detected, as suggested by Tomarkin, Davidson and Henriques (1990) in line with Russells' circumplex model, the approach and avoidance model became a methodological template due to its relation to the emotions primal levels and context. Edmund T Rolls (2001) provides a neuroscientist's anatomical based account, which suggests that the emotions are elicited in a goal orientated appraisal system of approach or avoidance within a reinforcement system of reward or punishment (see Fig 3.8).

Similarity of this account to Damasio's; 'emotions as homeostatic regulators' within his 'somatic marker hypothesis' may be made. A brief historical psychology survey of the approach-avoidance model is presented by Elliot and Covington (2001), which sites the landmarks in its theoretical development.

However, the anomaly, which prevents the universal implementation of the approach –avoidance model as an economical detection method for the range of the emotional palette, is the state of anger. As Eddie Harmon-Jones & John Allen (1998) demonstrate, the emotion anger shows approach tendencies whilst pertaining negative valence, thus sharing the same dimensional space as joy which has approach tendencies with positive valence. Yet their study confirms the working model of the approach-avoidance system, and that understanding of further dimensions of detection may be required for an all-consuming system.

Finally, it should be noted that there are two distinct directions in Asymmetric Hemispheric deciphering of emotional Valence levels. As described above one is the detection of transitional states between modes of emotional being. This animated system is built upon an underlying baseline of hemispheric asymmetry, which in part is claimed to account for the propensity of participants to respond in particular manner.

Thus Valence reactions to stimulus may be predicted as Tomarkin et. al (1990) demonstrate, whereby participants with right hemispheric weighted baseline responded with greater intensity to negative film clips than those with a greater left hemispheric baseline. In a later study using female participants Robert Wheeler, Davidson, and Tomarkin (1993) were able to predict greater valence reactions to film clips for both polar baseline resting biases. In a more recent study, Sutton and Davidson (2000) performed a similar test to affective words, again being able to gauge the relationship between baseline states and the ensuing related reaction.

Concurrently Mark Cavazza et.al (2014a, 2014b) is undertaking a series of experiment which utilise simultaneous fMRI scans alongside EEG at electrode position F3/F4. The fMRI scans confirmed pre-frontal BOLD asymmetry to verify the use of AHD. In a novel real-time neuro-feedback interactive system, empathy levels were derived via Valence levels to drive an unfolding animation narrative. The above studies demonstrate that AHD electroencephalography has the potential to unveil some of the underlying facets of emotion, or as Coan & Allen state EEG can serve to highlight the moderators and mediators of emotion and affect.

3.6 Support for EEG Arousal Detection.

Above we have mapped out a potential method with peer support for detecting the emotional dimension of Valence. Whilst the anomaly of distinguishing between joy and anger prevents a universal detection method, it may be that with further dimensions such as detection of arousal levels it may be possible to address this anomaly. In terms of detecting emotional Arousal via EEG there is not such a defined comparative body of work or formulated method. There are however a number of speculative studies, which explore potential sites, metrics and methods.

Aftanas et. al (2004) report that both anterior and posterior lobes are involved in both high and low Arousal signals compared to those of neutral signals in emotional activation. Using a 62-electrode channel configuration, they observed an increase of right hemispheric activity in posterior regions and increased activity in the left anterior hemisphere. They also noted synchronisation changes across all frequency bands of Delta, Theta, Alpha, Beta, and Gamma; yet do not distinguish these changes from Valence signals.

Kroupi, Yazdani, and Ebrahimi (2011) also tested how Arousal may register within different spectral bands associated with the brain states. They found that Arousal may be detected by Theta in the right frontal cortex, that the left dorsolateral prefrontal cortex is activated by Arousal in Alpha and Beta, further that the right Central lobe region and the left temporal region are positively correlated with Arousal in Beta band. Kroupi et. al (2011) state that the complexity of emotion is responsible for this. As emotion is influenced by contextual environments, situations and different structures it is acceptable to see different neural regions activated in different subjects regardless of whether they seem to be experiencing the same emotion. This is quite significant, for it implies difficulties in formulating a universal framework of measure.

Following Chopin's proposal that arousal excitation presents a higher Beta power and coherence in the parietal lobe, with simultaneously lower activity in Alpha, Bos (2006) tested Alpha and Beta power, and Alpha/Beta ratios in the frontal cortex using a minimal set up 3 electrodes for Arousal signatures. He reported that in a binary classification structure, F3/F4 Beta power and the singular (Fpz) Beta frequency were the most successful methods tested. The study used 5 participants with a very small data set, yet no subsequent studies were conducted. The best-obtained single performance rate for an individual's Arousal was recorded at 97.4 %.

Yoon and Chung (2011) found a preference for high Arousal detection using Beta and Gamma bands in a study that enlisted 90 International Affective Picture System (IAPS) images for elicitation. However their claims of a 90% confidence of affective arousal detection from feature vector combinations from the temporal, central and occipital lobes, are based on a single participant with no follow up studies.

Thus we may view; neural measures of Arousal are considerably more elusive and are reported to be dispersed across multiple frequency bands and electrode locations. As Kroupi et.al (2011) suggest it maybe individual contextual differences that account for this.

Chanel et. al (2005) undertook what is increasingly becoming a prevalent method, of a multimodal speculative approach. They considered the Arousal dimension using EEG, GSR Plethysmograph, Respiration, and temperature sensors. Using images with predefined emotional keyword tags and value ratings from the IAPS database in a binary classification structure, they were only able to achieve rates marginally above random. This was for EEG, physiological sensors and their fusion. However, when they measured the signals against the participant's own ratings to the stimulus, they were able to obtain one individual instance of a successful classification rate of 72%. The other three participants still oscillated around the random level. This reporting of higher variability for Arousal responses than for Valence is consistent within EEG studies. It may be simpler to report whether one feels more positive or negative, than the level of Arousal one feel. This may be due to a relational or individual nature, where variability arises from its tight entwinements with previous encounters and perspectives, which Kroupi et. al (2011) expand due to the different structures. Further it may be influenced by the experimental design.

EEG feature	*Location area*	*Frequency band*
1	*[PT;P;O]*	*$\theta 1$ (4-6Hz)*
2	*[PT;P;O]*	*$\theta 2$ (6-8Hz)*
3	*[PT;P;O]*	*γ (30-45Hz)*
4	*[AT;F]*	*$\alpha 2$ (10-12Hz)*
5	*[AT;F;C]*	*$\beta 1$ (12-18Hz)*
6	*[C;PT;P;O]*	*$\beta 3$ (22-30Hz)*

Table 1: EEG electrodes and corresponding bands.

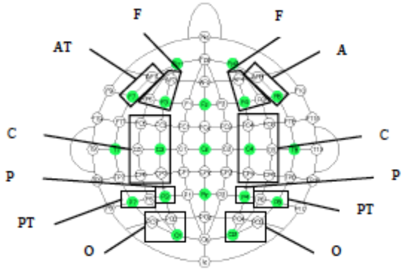


Fig. 1: Top head view with EEG electrode locations.

Figure 3.9. An illustration of Chanel's electrode configuration and feature vectors used to determine Arousal detection (Chanel et. al 2005).

For their experiments, Chanel et. al (2005), extracted 6 feature vectors, where only 1 feature covered the Pre-frontal cortex, with the majority measuring parietal regions (see Fig 3.9). Nearly all of their feature vectors incorporate the Occipital lobe. Naturally signal deviations will occur in this area when any visual stimuli is presented and it may be that activity and activation are being somewhat confused. Unfortunately there is no reporting for individual feature vectors, which may have given a clearer indication of this. If signatures for pupil dilation, and other visual processing (such as light, colour, movement) could mapped, and extracted from the EEG signal it may leave a clearer indication. Alternatively pupil dilation signatures in the EEG could be extracted as feature vectors for classification. Regardless this method does set the trend for modal fusions between neural and physiological sensors in attempt to provide more robust results.

Koelstra et. al (2010) in a multi-modal approach tested Valence and Arousal level responses to music videos with a population of 6 participants. For EEG they used 32 electrodes and their lateralization to report a decrease of right posterior Alpha power for higher states of Arousal. In single trails using binary classification, they were able to achieve a 55.7% rate average for Arousal with a maximum individual rating of 67%. These were on par with their physiological sensor signals of 58.9% successful classification accuracy. It maybe important however to note that with this method their

Valence detection levels falls much lower than peer reported levels to only just above the average random level.

Koelstra (2012), expanded on this in a subsequent study using 32 participants to create a database containing their recorded spontaneous emotion signals, for others to test potential algorithms and automatic classifiers. Again they were able to support their thesis of negative correlations in the theta, alpha, and gamma band, and importantly central alpha power decreases for higher arousal matches. Whilst they found high variability in subjects Arousal levels and reports, through their multimodal approach they were able to procure modest increases for Arousal.

Khalili and Moradi's (2008), multi-modal approach was able to achieve higher classification rates via EEG, than for physiological signals and the fusion of the two. In a 3 state Valance and Arousal classification structure of Positive-Excite/ Negative-Excite/ Calm, they achieved 51% classification rates, When the same data was condensed into binary class system, this rose to 65%. A common statement found in papers speculating on EEG Arousal detection is for authors to indicate that there is the potential for detecting Arousal in the EEG signal, without a specific determination.

Soleymani , Pantic and Pun (2012) used 20 objectively selected films clips and associated ground truth emotional keyword tags, and median Valence and Arousal scores acquired via a SAM test to establish tertiary classes for Arousal of ; Calm, Medium Aroused, and Activated. Using EEG, pupil dilation and Gaze distance, they were able to achieve 76.4% successful classification rates for Arousal. When the single modality of EEG was considered this fell to 62.1 %. In both results the random indicator was 33%, but it should be noted that only selective participants were included in their results. They also make a strong case for potentially further improving these results through additional modalities such as facial expression.

In the creation of their MAHNOB-HCI database Soleymani, Lichtenauer, Pun and Pantic (2012) were the first to precisely synchronise the five modalities of eye gaze, video, audio, peripheral and Central Nervous System in an emotional response scenario. This included 27 participants watching 20 short film clips. In their 3 class Arousal structure of; low Arousal, medium Arousal, and Activated, for EEG they were able to achieve 52.4% successful classification rates, whilst the fusion between EEG and gaze data increased this to 67.7%. In both the above experiments they state that good user independent arousal levels can be achieved. An important aspect they raise which has connotations for all emotion experiments, is for further researchers using their database to carefully consider participant recruitment as this can make a big difference in the results. A participant motivated by the experiment and the right knowledge for filling in the questionnaire is desired, as a participant attracted by monetary compensation may not have the desired motivation, communication skills or awareness of articulating their emotions in line with the questionnaire.

Huang et. al propose a technique entitled Asymmetry Spatial Pattern (ASP) for extracting features related to Valence and Arousal. Whilst they claim error rates of only 17.54%, they clearly state they selectively discarded data that they subjectively felt did not match ground truth labels. In their proposed ASP technique to extract features for EEG-based emotion recognition algorithms they employed K-Nearest Neighbour (K-NN), naive Bayes (NB), and support vector machine (SVM) for emotion classification. Whilst the average accuracy rates for Valence and Arousal are reported at 66.05% and 82.46%, respectively, it would be important to understand the data they selectively discarded. Huang et. al, experiment highlights an important issue where the focus is removed from the understanding of Arousal per se, to place emphasis on tackling emotion experiments as an engineering issue.

This is made explicit by AlZoubu, Calvo and Steven (2009), who state that their intentions are not to focus on the neuroscience behind affect nor to speculate or its implications, rather to focus on building automatic classification systems that may be potentially used for EEG emotion detection.

Jirayucharoensak, Pan-Ngum., and Israsena (2014), declare an EEG-based emotion recognition system "implemented with a deep learning network and then enhanced with covariate shift adaptation of the principal components". Through this the highest rates were able to achieve 52.05% successful classification on three levels of Arousal states. They state that their DLN proposal outperforms SVM and Naive Bayes classifier. They site inter-subject differences as one of the main issues to address.

Thus whilst this interest and direction of resources to and from the Affective Computing field is beneficial, it may be noted that there is currently an over-emphasis on tackling emotion research from solely a engineering perspective. This is with a view for the creation of automatic classification systems, which does not address or further explore unveiling the underlying measures of Arousal.

Horlings (2008) presents an acute example of how such proposed systems may not function as intended when moving from test data to the real world, and controlled condition experiments. Whilst in their system protocols they were able to achieve good binary distinction with training data, when implemented with real data, their classification rates fell dramatically to 30-40% correct classification for both Valence and Arousal. For Arousal this rose to a single high individual classification of 70 %, and further increased to 80% successful classification, where data for classification was highly selective; in that only data which was considered at extremely high or excessively low readings was selected.

In a study exploring the relationship between Arousal and Attention in simultaneous fMRI and low frequency EEG (5-9 Hz) recordings, Foucher et.al (2004) suggest a positive correlation between them, where higher arousal activates higher dorsal- lateral prefrontal and parietal cortex regions. Importantly, Foucher highlights the temporality of Arousal levels that need to be given further consideration. They state that Arousal may function in smaller time frames than is measured by a single averaging over a whole epoch of stimulus. Arousal transitions may occur with high variability, both in power and time frames that may be due to personal associations and definitions to phenomena out in the world. Thus, if emotion arousals are linked to previous episodes, then it may be why this indicator has seemingly more elasticity to phenomena. Surveying a scene of multiple objects regardless of the presentation format may elicit different Arousal responses and time frames, thus real time systems this may need to cater for this is any interface, classification protocol or experimental design.

In a study eliciting emotions via memory recall with 10 participants, Chanel et. al (2009), assessed Valence and Arousal indicators in short time spans to attempt to distinguish between the three states of: negatively excited, positively excited, and calm-neutral. For Emotion detection for their 3 classes they were able to achieve an average of 63% successful classification rate. When this was reduced to 2 classes this increased to 73%. When considering single class, of neutral-positive in a binary classification they achieved 96% accuracy for Arousal. Thus we can see whilst there are many proposal which indicate different methods for obtaining Arousal measures via EGG, there is not a definitive or consensually agreed method.

3.6 Conclusion.

In this chapter we have considered the potential of EEG for the detection of emotion using the dimensional framework. Whilst the peer review suggests support for

Valence detection using the AHD approach-avoid model, no such consensual method could be found to support Arousal Detection via EEG; instead we found a more speculative approach, without conclusion.

It is important to consider how this bears upon our original intentions of wanting to assess emotional responses to cultural artefacts in natural environments. Both the theoretical frameworks and supported detection methods for EEG imply, rather than a transparent window to the emotions, the type of access we may consider is both limited and general. The dimension theoretical framework whilst allowing some form of temporal access is limited only to Valence and Arousal levels from which we may infer a relative state. For EEG, the review only supports detection for the Valence vector, and as such this is using the approach-avoidance model. Within this there is the anomaly of not being able to distinguish between Joy and Anger to prevent a complete universal method. Whilst we may still map our approach-avoid model on to the Valence dimension we can in no way view this one-to-one relationship with Russell's correlative keywords.

Thus we may enlist the AHD model for the Valence dimension, but for Arousal we will have to use a speculative approach. So it may be important to again consider Berger's original thesis that Alpha is a signifier of neural activity, and consider whether excited and calm states as discerned by Alpha spectral power may be mapped on to the Arousal vector through correlations to the participants self reports. Through a process of assessing a single electrode pair site and frequency band we may be able to start a systematic approach towards either confirming or eliminating this possibility, and also test any differences between real and controlled experimental conditions.

We also noted that all EEG studies we reviewed are located within laboratory settings and that no studies match the type of contexts in which we hope to conduct our studies. If we return to our intentions of wanting to assess responses to cultural output in their natural environments, we find we are iteratively having to curtail our expectations of making these responses transparent and having to accept that these experiences will remain largely opaque, and any inferences we highlight will be of a very general and low resolution nature. However the full literature review suggests that the whilst the field of emotion research is complex and perhaps problematic it warrants study, and it is by engaging with these complexities and new contexts that possible further ground may be covered, even if only to eliminate possible directions.

CHAPTER 4

Experimental settings

4.1 Introduction.

An important task, given the goals of this thesis, is to try to bridge between laboratory based studies of emotion using EEG and the 'in the wild' situations of emotional responses that we hope they bear on. Natural observation is the study methodology of observing phenomena in its natural environment. It is a method traditionally used in the behavioural sciences. Aptly, in the context of emotion research, Charles Darwin (1934) used this form of methodology for his field studies from which emerged his particular theories of natural evolution and species wide similarity of emotional expression, to form the foundations for the discrete theory of emotion.

Whilst modern emotion research has largely resided within laboratory settings, there is an increasing call for observing the natural occurrence of emotions in ‘in the wild’ studies which may allow for the field to gain further understanding in its investigation.

In terms of EEG emotion studies, experiments have traditionally been confined to laboratory conditions due to their reliance on clinical technology, and also to procure confidence in results that may arrive within the reliability of controlled settings. Recently a new generation of commercial EEG technology have become widely available. These low-cost, wireless and mobile headsets present the potential for new forms of experimental design. Thus in the following passages, in considering this movement from laboratory to real world scenarios it is profitable to look widely at the field of emotion research beyond EEG for examples and possible access points we can bring to this study.

4.2 Calls for Natural Emotion.

Both Jerome Kagan and Rosalind Picard profess in their writing the need for emotion to be considered in its natural state of elicitation. Picard's thesis 'Affective Computing' (1997) envisages advanced forms of Human-Computer-Interaction whereby technology interacts with humans in an intelligent manner, through an ability to read and respond to human affective states. Picard highlights that the very nature of the laboratory setting is not conducive to researching this goal. The laboratory experiment is often confined to small time frames and representational conditions that serve to affect any response given. Thus, whilst these responses may give us a window to the emotions, they may not necessarily align to the natural occurrence of emotions. Picard also cites that the artificiality of the laboratory type setting is further compounded by the tight demographic of participants that may access such environments.

Jerome Kagan (2007) empathically highlights the need for the inclusion of ‘Origin’ into the experimental design of emotion investigations. Kagan cites Origin as being an extensive form of continuity that intertwines local/personal narratives and cultural beliefs as inherent within emotion. For Kagan, the issue of subjectivity is central, and for knowledge to widen, steps should be taken to exploring emotion in its natural states and environments of occurrence.

Paul Griffiths & Andrea Scarantino (2009) share this viewpoint, for the true complexity of emotion to be unravelled and understood there needs to be consideration of its non-reductive factors. In their Situationist perspective on emotion, they adopt a stance that links and embeds emotion within its social contexts; which are scaffolded within the environment. Importantly Griffiths & Scarantino show that emotion in social situations is not a singular phenomena or necessarily a linear process, rather that emotions are adaptive negotiations dependent on unfolding actions within the environment (others). Thus, emotions do not function in the everyday or within interaction as they may be researched in laboratory strategies.

In experiments by Ian Davies and colleagues (Davies & Robinson, 2011) the important distinction is made of differential levels of emotional investment between real world and constructed settings. Whilst they find that within laboratory tasks, levels of emotional investment can often be found to be lacking, they remain equally sceptical of data-gathering techniques in natural settings in regards to repeatability.

Figure 4.1: Davies emotional investment experiments explored the impact of settings on results. Participants engaged with a computer simulation in a traditional laboratory setting (Left and middle), and also within a hybrid setting controlling a remote controlled helicopter via a live feed. (Right). (Davies, 2011).

Their response to this critical position is to consider hybrid 'Real World-Laboratory' conditions. In a comparative study demonstrative of this, participants engaged in a virtual racing game, followed by a remote controlled helicopter task, whose actions they observed and controlled via screens in the testing laboratory (see Fig 4.1). They surmised, in the latter task participants were able to engage in a choice and real consequence task, which resulted in higher investment levels. Simultaneously, they could be monitored in a repeatable method, thus meeting both criteria. Whilst it could be argued that the setting would still influence the outcome, it shows how it may be possible to move towards meeting both criteria of naturalism and repeatability.

Thus whilst there is much validity in the calls for widening the contextual and practical envelopes for emotion research, due care has to be given so that experiments and their results are robust, accountable and reproducible. Whilst it may not be possible to immediately address all the expansive calls, an iterative procedure of testing scenarios and technological detection methods may pave a way towards it. The single modality of EEG for emotion detection that we are focusing on in this project, will naturally not unveil the expanse of any emotions, but our findings may contribute to insights into one aspect, which can then be built upon. By carefully catering for some natural limitations in natural environments we may also be able to conduct experiments that meet both our aims of natural responses and also having the assurance that our measurements are solely the responses we are hoping to detect.

4.3 Natural Settings Limitations: Technological Dependency.

For emotion studies to be conducted ‘in the wild’, there is a need for an appropriate technology that meets a basic criteria. Any data harvesting systems would need to be lightweight, robust, portable and unobtrusive. It would need power capabilities that allow it to run continuously over large time frames so it could be used in longitudinal studies. Finally, it would need to store and make data available flexibly. In this way the device would be rendered invisible to the participants through habituation and may serve to obtain the truest streams of natural data.

Across the field of emotion detection many commercial and medical systems are capable of harvesting physiological signals for inferring emotion. Whilst traditionally these were designed for stationary recording procedures, a new trend for mobility has emerged.

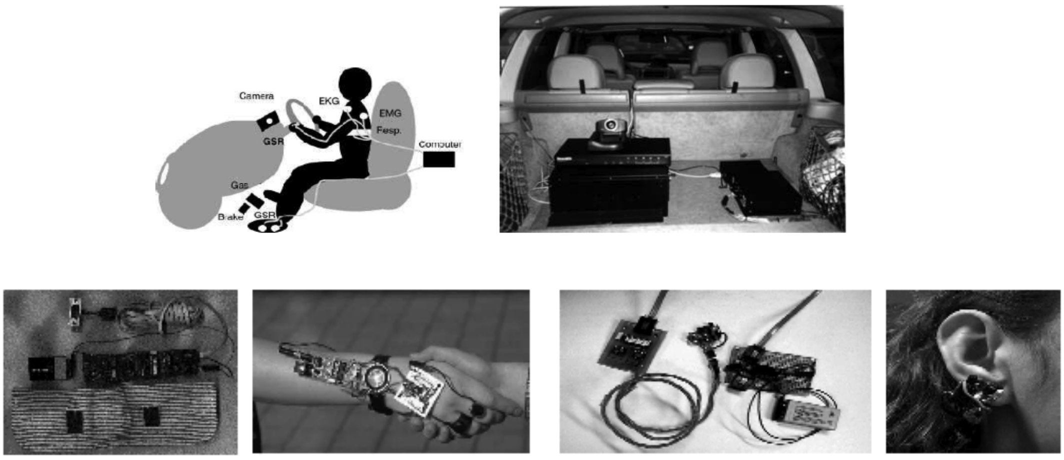


Figure 4.2: Healey used mobile multi-modal sensors for her experiments into naturalistic emotion recognition and classification. (Top row) The sensors configuration for the driving task, (bottom row) some of the prototype sensors she constructed for the task. (Healey, 2005).

In 2000, as part of her PhD thesis Jennifer Healey (2005) conducted investigations into developing an appropriate mobile technology for naturalistic emotion recognition and classification. Her experiment entailed participants driving a car through variable traffic conditions whilst wearing a number of sensors that recorded their stress responses. (see Fig 4.2 top row).

As is visible (Fig 4.2 bottom row) with her assembled prototype wearable multi-sensor network for gathering the bodies signals, whilst they are adequate for obtaining ambulatory style data, they are obtrusive and cumbersome, especially in comparison to the standards and expectations we have of sensors in our contemporary moment.

Figure 4.3: Peter’s mobile wearable data glove. (Peter et. al, 2005)

In 2005 Christian Peter and colleges (Peter, Ebert & Beikirch, 2005) also developed a prototype technology to meet the needs of sensing physiological data in the wild. Their wearable technology consisted of a glove (see Fig 4.3), which collected Galvanic Skin Response (GSR) and Heart Rate Variability (HRV) data to be transmitted wirelessly to a mobile base unit. Their system was cable of running continuously for a week before any battery replacement or recharge. Whilst their data glove presents the potential for collecting good data, its bulky, restrictive and visible design must be noted, and may potentially interfere with any formal experimental results.

Figure 4.4: Picards Galvactinator’s progressive transition to the Q sensor (left to right)

The Galvactinator co-produced by Rosalind Picard is a similar product (Picard & Scheirer, 2001). Here sensory measures of Electro Dermal activity (EDA) were made visible and available through a colourful Led Light in place on a type of data glove. In experiments with a conference audience, Picard was visually able to note en masse, audience arousal levels in response to specific activities. The succinctly stepped development of this idea led to the realisation of the Q sensor produced by Affectiva (see Fig 4.4) (Picard, white paper & Poh, Swenson & Picard, 2010). Here the wearable device is in the form of a wristband. The uniqueness in these devices is the relocation of the sensing electrodes to a non-intrusive location on the wrist, which aids it invisibility and interference.

The friendly usability of the Q sensor is an element of its appeal, the sensor automatically takes readings of EDA, motion and temperature when worn, which is then logged and displayed via its specialised software and in this way eradicates any need for base stations or worn smart devices.

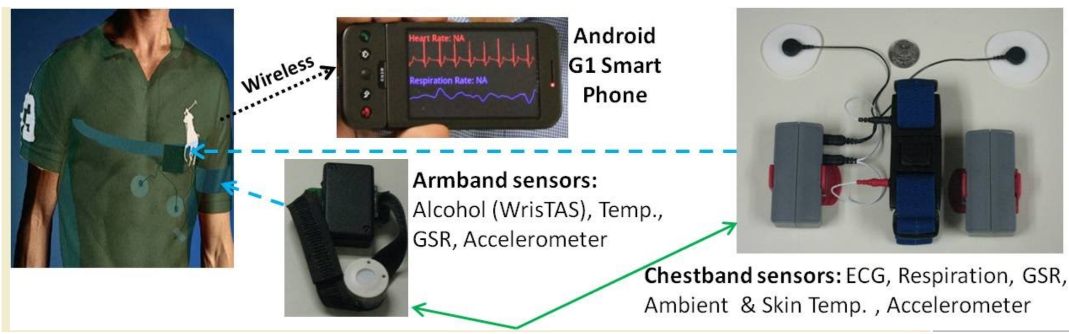


Figure 4.5: The Autosense multi-sensor system. (Ertin et. al 2011).

Taking advantage of the omnipresence of current smartphone devices, The Autosense system uses multiple sensors of; ECG, RIP & GSR, to obtain HRV, breath rate and skin conductance levels. In a philosophy of more sensors means more data means more information (see Fig 4.5), the data is transmitted for viewing and recording via a Smartphone device (Ertin et. al, 2011).

Each of these directions of minimal or maximal sensor data gathering has their particular merits and downfalls but together they cater for a range of experiments that allow affective data signals to be gathered beyond the laboratory space. It is perhaps by looking forward towards emergent technologies we can sense how the protocol and popularity of ‘affect in the wild’ experimentation may evolve, and thus the value of conducting and exploring research methodologies on this topic in the now.

4.4 Beyond the Laboratory: Emerging and future sensor technology.

Taking into account the wide field of physiological sensors we can begin to envisage the potential of how natural observation experiments may function to meet the needs of the natural emotion researcher. The health and fitness industry is one component of medical research driving the development of miniaturised technology. This technology is capable of gathering physiological readings of a users state and then feeding that information back to the wearer or the assessor in ways that can be meaningful either diagnostically or personally.

Figure 4.6: The Basis Smartwatch [16].

If we consider the prototype technology Healey produced for her thesis, we can consider it cumbersomeness against technology of similar capabilities emerging or verging to emerge little more that a decade later.

Smartphones can be aligned to mini-computers that have proliferated into our communication devices. It can also be seen that this continual miniaturisation is leading to the development of further devices; Smartwatches. The Basis Smartwatch (Health tracker) available for commercial release in 2013 is one example of such a device being engineered to present physiological data to the user (see Fig 4.6). This aspect of the device has a number of built in sensors for detecting heart rate, perspiration levels, body temperature and motion. Whilst The Basis is reminiscent of previous ‘sports’ monitoring technology and Smartwatches configured for communications, this device offers continual basic tracking data that classifies activity for up to four days before a

[16] Retrieved from https://store.mybasis.com/media/catalog/product/p/e/peak_time_ black_three-quarter .png

battery recharge. Within the increasing complexity of this technology it is becoming a standard to incorporate API's, and Software development kits (SDK), which promise accessibility and customisation for any particular need. The ability to link to smartphone devices only increases their power to allow data harvesting in an invisible manner. This generates the potential for limitless possibilities of configurations of experimental design.

Importantly, this interest and practicality of the roaming capabilities of sensor technology is also paralleled in the field of electroencephalography. Traditionally, EEG laboratory setups comprised of a laborious process of gluing and wiring electrodes to a participants scalp. This lengthy and sometimes cumbersome process also creates a restriction of natural movement for participants and situates the experiment.

Figure 4.7: A new generation of low-cost mobile EEG technology have emerged marketed at gamers, creative's, researchers and developers. (Left to right) Imec Hoist [17], Emotiv Epoch [18] ,Neurosky Mindwave [19], InteraXon Muse [20].

A new generation of commercially available EEG headsets now offer wireless forms of data transmission that open up the potential for new forms of naturalistic experimental design. Interaxon, Neurosky, Imec and Emotiv are amongst a handful of organizations that are popularising and leading the field of portable EEG headsets. Each of the new EEG headsets is lightweight and simple to fit, they are simply placed on the scalp. Each has its own variation of the standardised 10-20 system electrode configuration (see Fig 4.7). These range from a singular, to fourteen sensors arranged over key areas of the scalp. Thus, whilst they do offer limitations in comparison to clinical EEG in terms of the amount of data sites which can return signals, this may be weighted against the freedom they offer as to the range of stimuli and environments

[17] Retrieved from http://www2.imec.be/content/user/Image/Staalhemel.jpg
[18] Retrieved from https://emotiv.com/bitrix/components/epanel/store.headset /templates/.default /images/head.png
[19] Retrieved from http://cdni.wired.co.uk/620x413/g_j/headset.jpg
[20] Retrieved from http://nvate.com/wp-content/uploads/2013/06/1-e1372742855935.png

they allow for experiments, especially in less formal conditions.

The headset selected for this project is the Emotiv Epoch Education Edition, which provides full access to raw EEG data, alongside an accompanying SDK. This meets the requirements for our investigation of an economical, robust, portable headset where we can explore detecting natural aesthetic-emotional responses.

Further generations of these headsets are emerging, which respond to en mass usage issues fed back to the organisations. For example, Emotiv are in the process of releasing a crowd-funded 5 sensor set up, which eliminates the requirements of re-moisturisation of the felt contact pads of the Epoch headset, which is the main issue in limiting extended continuous timeframes.

As with other technology these reported requirements will drive further developments of miniaturisation, optional algorithms (such as automatic artefact reduction) and yet to be considered forms of potential brain mapping.

Figure 4.8. Imec's thermal powered EEG headset [21].

At the forefront of such future developments are Imec, an organisation that directs its research activities towards the production and development of Nano technology. Building on the previous success of their wireless EEG Headset they are now exploring how such devices can be constructed to run on thermal means to eliminate the need for batteries or other external power sources (Leonov, 2008) (see Fig 4.8). Here the body's thermal emissions provide the power source for the headset, with the consequence of a continuous stream of data that is relative to the users own power timeframe.

[21] Retrieved from https://tomography.files.wordpress.com/2007/11/resized-wearable_press.jpg?w=460

Figure 4.9: MC10 's wearable sensor technology [22].

This notion of a self powered miniaturized sensor technology is also shared by American company MC10, who have developed miniature ‘elastic electronics’ that can be applied directly to the skin in a form that is analogous to a medicinal plaster (see Fig 4.9) (Kim, Lu, Ghaffar & Rodgers, 2012). This is capable of transmitting self-powered continuous data wirelessly and is a glimpse of how invisible non-invasive sensors may operate in the near future. It is not difficult to imagine how a miniature EEG sensor network placed on the scalp may function within this trajectory.

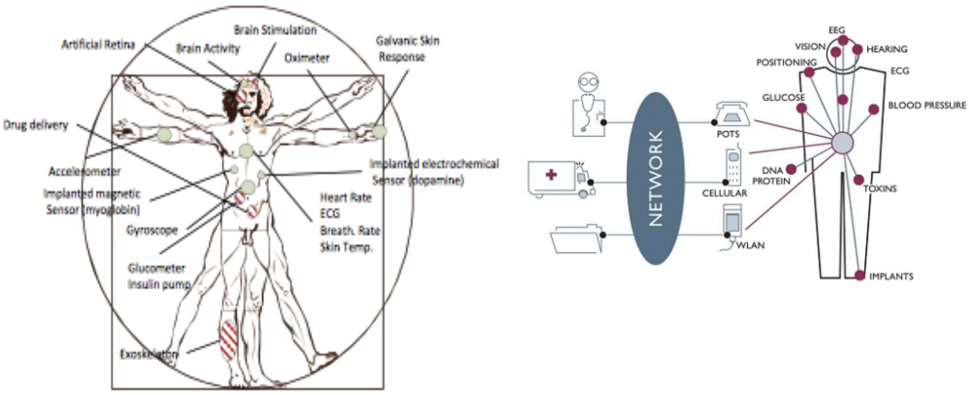


Figure 4.10: Body area Network are range of physiological sensors [23].

With such forms of invisible wearable sensors, the projections of Body Area Networks, (Brandao, 2012, Chen, Gonzales, Vasilakos, Cao & Leung, 2012) whereby the body holds an array of sensors that continuously transmit variant data becomes more plausible (see Fig 4.10). Thus further progression and development towards this

[22] Retrieved from http://www.mc10inc.com/wpcontent/uploads/2013/12/MC10_ Biostamp.jpg

[23] Retrieved from http://www.wifinotes.com/computer-networks/body-area-network.jpg

network present the likelihood that emotion and EEG research will find it easier to gain acceptance for out in the world experimentation.

Above we have taken into account a range of mobile sensors, and sensor set-ups, which may contribute towards detecting emotion in natural settings. In this research project, the particular focus is on a detailed exploration of what we may neurally be able to detect about the emotions via EEG as a standalone technology. Through a comparison of laboratory and 'in the wild' settings we may be able to gain an understanding of the potentials and limitations of this current cutting edge technology. Further through this assessment we can consider its validity for use in more complex multi-modal sensor arrangements, which may provide further richer possibilities of detailing of emotional responses.

4.5 Self Report : Data Labelling and Classification.

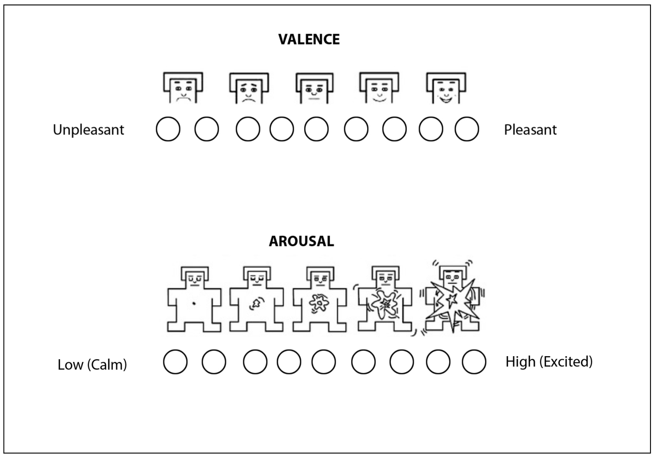


Figure 4.11: A 2 vector 9-point Self Assessment Manikin (SAM) Test.

Thus far, this section has considered the technological acquisition of physiological data signals in naturalistic settings. Another key issue for this method is the consideration of how such acquired data can be made meaningful through its quantification and classification.

Jennifer Healey (2011) writes that the biggest challenge facing affective computing is the accurate labelling of affective data. To date the most viable form of understanding and labelling a subject's emotional state to their signals is through forms of self-report. In the controlled laboratory, post experiment surveys, questionnaires, interviews or Self Assessment Manikin (SAM) (see Fig 4.11) tests are normally taken at the researchers discretion to ensure that the sensor data and reports are relationally time-stamped, and reported back meaningfully. For a SAM test the participant simply places a marker to signify their feeling in relation to the given dimension. In real world situations however, this becomes more problematic.

Here the subject or participant is continuously undertaking tasks away from a guide or researcher and this creates a higher potential for the mislabelling of data which can arise when there is a delay between the experienced emotion and its detailing.

Healey notes that labelling is always biased by the current experienced emotion, therefore delayed labelling after an event has a larger potential to be affected. As information of an emotional experience occurring within a context is translated from short to long-term memory it may become considered within a current or different context. An emotion labelled as it occurs may be defined within its felt intensity. A previous emotion reflected upon, may cause its reporting to be up or downgraded in the scale of its intensity and nature due to its comparison to other previous memory's in the subjects history/ consciousness. In a similar manner to the recalling of a dream, its intensity and detail is lost with time, and only the peak or immediate definitions made post-dreaming stay in memory.

Another potential flaw of self-report is the variant levels of emotional awareness and emotional intelligence of participants. Within cultures, where the focus is traditionally on logic and rationalisation over intuitiveness there may be less awareness, or ability to articulate this form of felt intelligence. In his informal pocketbook on EQ, Rob Yeung (2006) demonstrates that the first stage of becoming emotionally intelligent is to become self-aware. The greater the self-awareness the greater the propensity one has to note feelings or name and label emotions as they occur. Perhaps for natural emotion research some form of pre-experiment familiarity with EQ may be an enhancing addition, to receiving more detailed alignment between signals and surveys.

Regardless of these factors, self-reporting still remains the best means of accessing the private spaces of subjectivity and the labelling of subjective data. In her experiment Healey (2011) reconfigured the P-A-D model to produce smaller scalars in the range of 1-7 with the central node '4' relating a to normal feeling/state. Her motive here was to make the labelling simpler and clearer with a neutral default position for the participants. 13 participants took part in her study conducted over a 7-day period. The first two days were given over for acclimatisation to the method. Using GSR and ECG sensors, a physical activity monitor, and a mobile journal to mark timelines with the beginning and ending of an emotion, Healey concluded that this form of triangulation of data leads to better ground truth labelling.

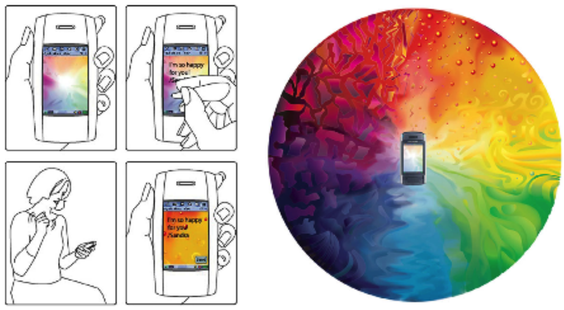


Figure 4.12: Kristina Höök's eMoto emotional text messaging interface. (Höök, 2009).

Design is another discipline, which has an interest in researching the natural occurrence of emotion. Kristina Höök's practice engages in a design-centric approach to exploring continuous and spontaneous affective experiences ' in wild' settings (Höök, 2009). Höök views emotion as arising within interaction, and clarifies this through her focus on the dialectic of embodiment, which she modifies to the term 'affective loop experience'. Here the interpretation (labelling) is solely with the user rather than the system. Her position is a recognition that emotions are co-created and inseparable from aspects of life. Höök develops devices which consider 'affect' both through forms of communication and potential feedback devices that have affective 'in the wild' annotation properties.

eMoto (see Fig 4.12) is a form of emotional text messaging interface, where physical gestures such as shaking the phone allows a choice of colours and animations to be constructed in the background of the text message. These variant colours and animations are believed to provide the receiver with a sense of the sender's emotional state and tone. Naturally through usage this becomes a consensual learned language of communication.

Figure 4.13: Höök's Affective Diary. (Höök, 2009).

Höök's affective diary (see Fig 4.13) is a digital journaling format technology, which allows users to reflect and annotate their emotional states. An armband sensor collects arousal (GSR) and motion data that is then transmitted and logged onto a visual timeline. The data is displayed in a simple representational human like form and changes in shape and hue in relation to the sensor data. At any time the user may digitally scribble annotations onto the timeline linking it to the current image to create a reflective platform for their emotional patterns.

In a month long study with 4 participants this form of feedback proved useful in both the logging and annotating of data, and importantly allowing the user a sense of familiarity, and a gauge to their inner emotional world and the variant triggers and responses.

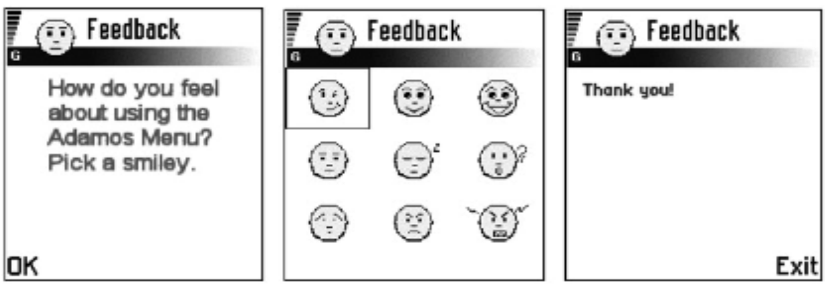


Figure 4.14: Minna Isomursu and colleagues, emoticon annotation interface. (Isomursu et. al 2007).

Another aspect of Design that calls for emotional evaluation is the assessment of human interaction with products. Minna Isomursu and colleagues (Isomursu,Tahti, Vainamö and Kuutti, 2007) conducted a comparative study of 5 self-reporting methods of emotional responses to mobile apparatus in natural type settings. Alongside SAM and emo-card testing they explored three further experimental methods: (i) the drawing of representational images following the interaction, (ii) a feedback emoticon program included on the device (see Fig 4.14), and (iii) an ‘Experience Clip’ which was a audio-visual recording device (a phone). Of their three proposed methods, they deemed the Experience clip the most successful for gathering annotated emotional data.

This method incorporated the participants exploring the ‘product’ in known pairs. Whilst one engaged with the device, the other was able to record ‘up close’ in a manner that did not impinge on the activity of the participant as may occur if the researcher was present. The researchers surmised for their task this provided the most useful information that alongside transcripts of dialogue could then undergo facial emotional expression and vocal recognition of the data.

These few examples highlight both the necessity and problems of annotation and labelling. It is essential that these issues be give due thought in any experimental design that attempts to explore emotion in natural settings. As with the above highlights each experiment needs to consider its context and it most suitable method, and also how meaningful the gathered data is.

In her text ‘For the people by the people’ Picard (2010) highlights this issue of the meaningfulness of data. She notes that data would be more meaningful if it was fed back to participants giving a form of heuristic value. If we consider her vision of affective computing and also how HCI devices function, it is mostly aimed at the individual rather than group level, and thus gathering and working with a single participant may allow a deeper understanding when multiple individual longitudinal studies are brought together for comparison.

4.6 Natural settings Location Tests

Drawing from the above findings, it became important to find an appropriate naturalistic EEG study method. Thus a series of pilot studies in the form of practical walk-through’s were conducted without participants to consider appropriate settings, stimulus, and annotation methods for EEG purposes. As no formal body of EEG research or experimental EEG method that we are aware of could be referenced for this, this process enlisted the art and design research method of 'reflection in action' and 'reflection on action'. Simply stated, this is the process of undergoing natural activity whilst maintaining reflective forms of objectivity. By conducting the pilot studies in this way, the experimental conditions could be experienced first hand. Practical issues could be identified and the potential nature and impact of elicited emotional responses could be subjectively gauged. The overall goal was to explore a wide range of stimulus from which the most suitable could be drawn. Considerations included sculpture, painting, animation, film, installation, light art, sound art, and live theatre. The particular stimulus selections detailed below were dependent on current exhibitions and also ease of gaining permission to access spaces. The selected cultural works used in these 'walk-throughs' were identified through web searches for exhibitions using artrabbitt.com, which is devoted to Arts listings. Further web searches were conducted for current theatre productions. None of the works in these pilot studies had been previously viewed, although there was an awareness of some of the artist oeuvre.

Each study iteratively incorporated findings from the previous study, and problematic issues addressed with potential solutions tested. The issues of consideration identified for the pilot studies were; (i) A real world setting, (ii) Repeatability of stimulus and setting, (iii) A stimulus of sufficient strength to elicit emotional responses, (iv) Appropriate annotation method, and (v) Control of variables. In each of the following instances these points were considered.

4.6.1 Experimental Setting | Example 1

Venue : The Serpentine Gallery. London
Artist : Jonas Mekas.
Format : Projected Artist Film/Video.

The Serpentine Galley is a compact popular public Art gallery located in Hyde Park, West London. It hosts a roster of international artists working across diverse mediums. Jonas Mekas is an artist who works mainly with moving image, making forms of video diaries that intend to celebrate joy. His imagery comprises of montaged snippets of personal hand held video footage; of places, peoples, and actions.

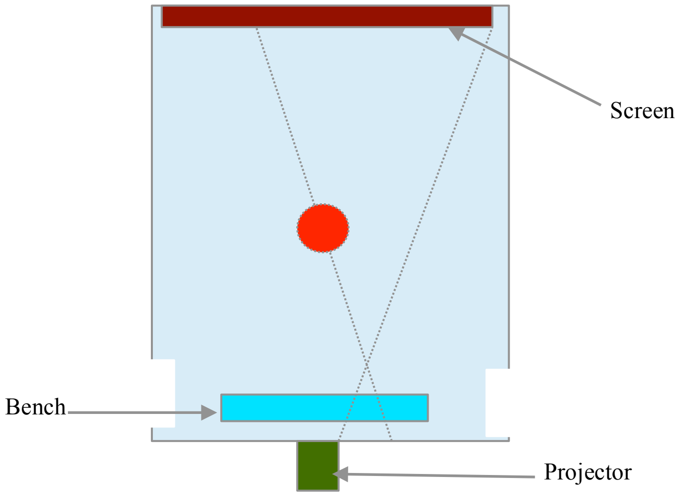


Figure 4.15 : Topography of the gallery space for pilot study: Example 1. (red dot= optimal recording position)

After fitting the headset, and wandering around the gallery, a few members of the public, who were curious about the headset, approached the researcher. This was an issue that would strongly impact on any formal experiment. Thus the most ideal space was sought in view to limiting this. The main gallery was a blackout space, it comprised of a full wall projection with a bench in parallel (see Fig 4.15). When attempting a recording seated on this provision, whilst there were no oral disturbances the viewing was affected by a constant rotation of people sitting down, getting up and walking past.

Thus a recording was made sitting on the floor, which whilst uncomfortable, presented an unimpeded view of the moving images. Here considerations were made of how to annotate one's sensations to the stimulus, a paper and pen were not possible due to the light levels in the space, and interruptions of the viewing, and the potential artefacts this may introduce into the signal. A possibility considered was how an inventory of facial artefacts such as blink signals, could be constructed as an annotation syntax, which could be read in the EEG signal. Whilst this may have potential in certain instances, it was felt that this would distract the viewer from their experience, to focus on the annotation protocol. Whilst such a setting and format of work had potential, the elicitation of strong emotions to the Artworks content was highly questionable.

4.6.2 Experimental Setting | Example 2

Venue : Tate Modern. London
Artist : Oskar Fischinger.
Format : Projected Abstract Animation

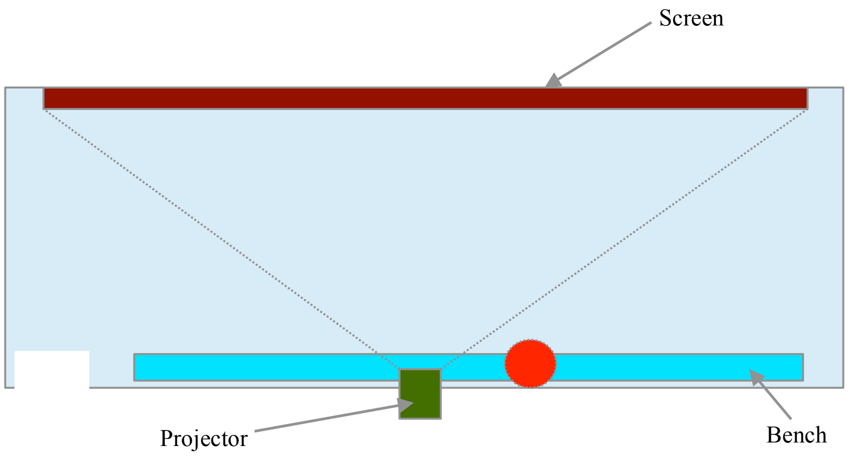


Figure 4.16: Topography of the gallery space for pilot study: Example 2 (red dot= optimal recording position)

After Permission was granted for the recording by the Tate office, all staff were made aware of the recording, and to offer any required assistance. This gave a feeling of safety to the researcher for wandering freely through the large building. A post-recording reflection on the Jonas Mekas exhibition concluded that a simple voice recorder would allow real time annotations to be taken, without disturbing the experience. This was implemented in this test recording.

The Tate Modern, is the most visited public modern art gallery in London, it houses both permanent and visiting exhibitions in a number of rooms arranged over numerous floors. Firstly, different works including hung paintings, free standing sculptures, prints, and moving images works were considered as potential future stimulus. As with the Serpentine the Tate is an extremely busy gallery. There was a consistent transit of the public throughout the spaces, such that when one attempts to view a work, it is often a shared experience with rotating others. Amidst the constant cacophony of random noise, there were a number of pleasant and polite inquiries as to my activities by others in the space. This again raised questions of repeatability of experiments and disturbances of this setting.

In spite of this a work by animator Oskar Fischinger's stood out as a potential setting. It is situated in a designated blackout space. Opposite a large wall projection of the work there is a bench for the audience to sit (see Fig 4.16). It is a work on permanent display and the short piece is looped. However, again whilst trying to attempt to make a clean signal recording a consistent level of movement and noise, served only to distract from the work, and hinder a clear recording. Further the nature of abstract visual works and their ability to illicit strong or prototypical emotions was questioned for this stage of the research.

In terms of the annotation method, a voice recorder was excellent. It allowed for a commentary on the work and feelings without distracting from the stimulus. It was concluded that for such works (inclusive of Jonas Mekas) it might be appropriate to only make comparative pre and post experience recordings, against a simple SAM test.

4.6.3 Experimental Setting | Example 3

Venue :The Hayward Gallery. London
Artist : The Light Show (Various artist)
Format : Light Installation works.

Figure 4.17: Works featured in The Light Show Exhibition included: Ollafur Ellisan (left), James Turrel (middle) , and Anthony Mcall (right) [24].

The light show was a survey exhibition held at the Hayward Gallery. It brought together a seminal collection of artworks that engage with the medium of light. This comprised of highly regarded International artists, including Anthony Mcall, James Turrel, Ollafur Ellisan, Jenny Holzer, Dan Flavin and Carlos Cruz-Diez amongst others (see Fig 4.17).

An EEG baseline was recorded in the Hayward Foyer. Once complete, both

[24] Retrieved from http://www.haywardlightshow.co.uk/

EEG and audio recorders were time-synced and started, and the researcher entered the exhibition. The arrangement of the exhibition was for the designation of either a single room or a large space for each work so that it may be viewed in a form of isolation. In turn, the researcher approached each work and spent a few minutes experiencing it. An annotation was made via voice recording, and then the next piece was visited. By using the audio recorder in this way it also became possible to locate in the continuous signal exactly where the viewing of any work began and ended, eliminating the need to continually start, stop, or place a marker in the Testbench recording interface.

Immediately noticeable was the quiet atmosphere and the availability of private space. At the Tate, Serpentine and other public gallery spaces visited it seemed chaotic, whilst here in a paying exhibition, firstly it was less crowded, and secondly everybody seemed to be more engaged in their own experiences. The only time the researcher engaged in any communication with others was when resting on a bench.

The headset was worn continuously for 1 hour 30 min's without too much discomfort. It was felt that this situation would be an appropriate setting for an experiment. It was an engaging stimulus, evoking contrasting experiences, and by association emotions. It was felt there was a reasonable potential for repeatability across participants, the annotation method was rich, informative and immediate, and the factor of different people in the space was nominal. Any disturbances of this kind could be further reduced through the careful selection of the experiment's timing.

Artwork Stimulus (in order of viewing)	Sequential Signal Movement	Movement from baseline mean
Baseline	==	n/a
Leo Villareal	+	+
Ceal foyer	-	-
Anthony Mcall	+	+
Doug wheeler	+	+
James Turrell	-	+
C Cruz-Diez	-	-
Conrad	+	+
Anne Veronica	+	+
Jenny Holzer	-	-
Olaffur Elliason	+	+
Baseline after		-

Table 4.1: This table show the Valence classification results for pilot study: Example 3. Column 1 lists the name of the artists work experienced in order of viewing, column 2 lists the sequential increase-decrease of the signal (mean value for whole stimulus) for each work viewed in order. Column 3 shows the signal (mean value for whole stimulus) in comparison to the baseline (mean). N.B (+) equals an increase, (-) equals a decrease.

The recorded data from this experiment was used to explore potential protocols for signal processing, classification and data handling. When comparing signal and annotation, simple relationships were noted; the most enjoyed work, gave the highest mean Valence value, whilst the least enjoyed gave the lowest mean Valence value (see Table 4.1). This assigning of the most and least and enjoyed work labels were based on my own subjective emotional responses to the works. Those which triggered positive pleasurable responses were annotated as being enjoyed, whilst the works which seemed to provide little or no response, or whose viewing I found no pleasure or interest in I annotated as not enjoyed. This is similar to how it was envisaged that the formal experiments would function.

It was felt that such a particular form of presentation with thoughtful planning, and the development of an on-person data capture (i.e. smartphone) would make an ideal experimental set up.

4.6.4 Experimental Setting | Example 4

Venue: Soho Theatre. London
Performance: Bitch Boxer by Chloe Jackson
Format: Solo Theatre Performance.

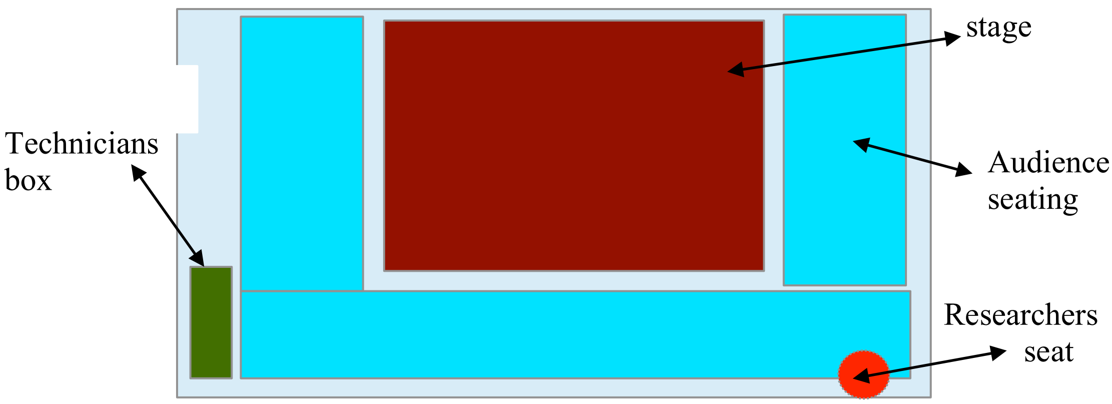


Figure 4.18: Topography of the Theatre space for pilot study: Example 4 (red dot= researchers recording position)

Bitch boxer is a solo theatrical performance written and performed by Chloe Jackson. It tells a story of a East London girl with a passion for boxing, who trains to fight and goes on to win a gold medal at the 2012 London Olympics. Through the narrative several personal relationships and themes in her life are explored. The Soho Theatre is an Independent theatre based in Soho, central London. It has a variety of performance spaces. Bitch boxer was performed in the upper theatre. Arranged around 3 sides of a demarcated square stage were approximately 100 seats organised in rows. The researcher sat at the end of an aisle parallel to the stage, one row from the back in a seat that was selected before the doors opened (see Fig 4.18),

Before the performance started the researcher recorded a short baseline. Due to the building being so busy and a lack of quiet spaces, it was recorded directly outside the performance space. Upon entering the theatre studio the EEG and audio recording devices were started. Post performance another short baseline was recorded.

Regarding the performance, it was immediately deemed highly appropriate for a participant study. Firstly as a mode of emotion elicitation; the performance was very engaging, it moved at a fast pace, through a variety of characters and situations, and through them, evoked a whole range of emotions. A talented single performer who held the audiences attention for its duration enhanced this. Due to the intimacy of the setting, the performer is in close proximity. Thus, the synchronised exaggerations of facial expression, vocal tone, gesture, and story telling which were further synchronised with the lighting, sound, and props, converge as modalities to fully heighten and exaggerate the intended expressed emotion. This increased the potential for variant clear elicitation.

Considering that all these aspect are repeated with as near precision as possible for every performance, it may be regarded as an excellent meeting point between 'real world' and laboratory conditions. One feels emotionally invested throughout the performance, and through this investment alongside engagement, is led on an emotional journey. The audience being seated also assists the recording process. This aids the experiment by reducing any potential movement artefacts that may distort or disguise the signal.

Post performance, walking through the area of Soho, the researcher listened back to an audio recording of performance whilst it was still fresh in mind, and noted it brought back much of the experience; feelings, memories of feelings, and further reminders of sections that had flashed past. Thus it was felt that despite a short duration had lapsed, the experience of the performance was fresh enough against which to provide some form of annotation. This could be further tested in the formal participant experiments. From this test recording it was confirmed that the theatre was the most suitable type of venue and stimulus for conducting a participant study in the context of this research projects aims.

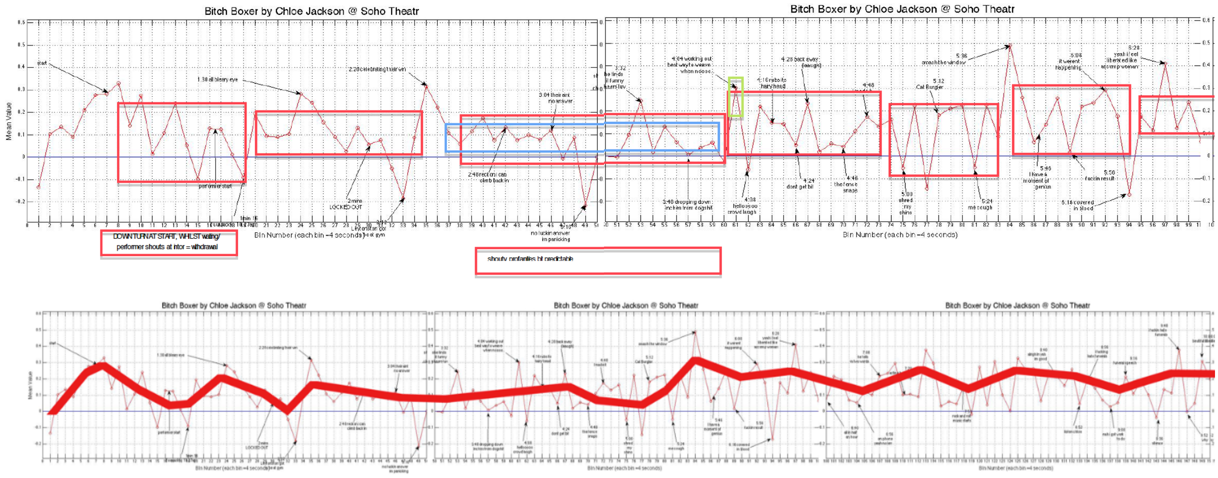


Figure 4.19: Speculative classification patterns were explored in the continuous data signal for pilot study: Experiment 4.

The recorded data from this experiment was used to explore potential protocols for signal processing, classification and data handling. A 20-minute section was transcribed and plotted, and patterns were noted with in the signal for potential methods of interpreting emotional movements within a temporal signal. This same section was also used to test Ocular Artefact (OA) removal, whereby the blinks were algorithmically removed from the signal (see chapter 5).

4.6.5 Experimental Setting | Example 5

Venue : The Lisson Gallery. London

Artist : Harron Mirza

Format : Audio-Visual Sculpture.

Two test recordings of temporal Contemporary Artworks were conducted at the Lisson Gallery on separate occasions. The Lisson Gallery is a highly regarded International private Art Gallery located in Central-North-West London, The gallery has a number of spaces arranged over two floors, which due to its private status are mostly quiet and free from others for the majority of the time.

The first test was conducted to Haroon Mirza's installation Pre-occupied waveform (2011) presented in the lower gallery. Mirza creates an assortment of audio-visual-sculptural type assemblages. These have various components that contribute towards a conjoined composition. The exhibited work can be experienced in different ways. One can statically survey the whole space, or circumnavigate individual aspects of the whole. The audio-visual elements can taken as a singularity, or individual units, further the audio topography can be isolated with eyes-closed. All propositions were tested.

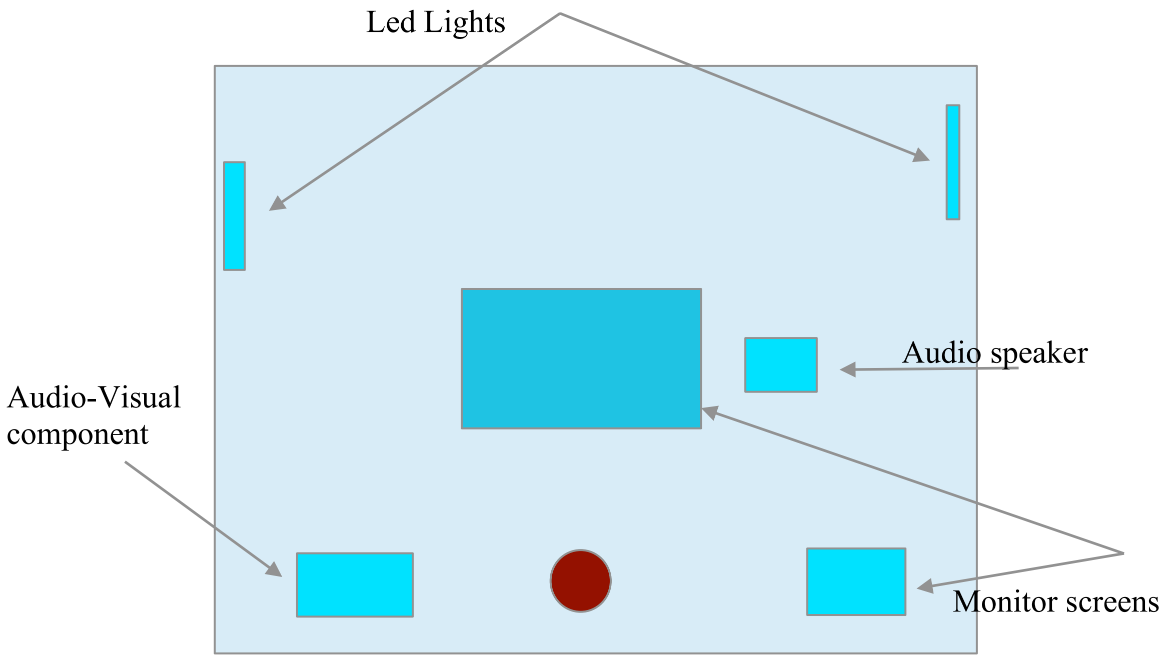


Figure 4.20: Topography of the gallery space for pilot study: Example 5 (red dot= optimal recording position)

In regards to these different forms of viewing potentials and the recorded signal; footsteps generate easily detectable artefacts in the signals, whilst head movements are recorded via the headsets in inbuilt accelerometer, thus such sections can be identified and extracted. Whilst the work may not bring to mind a strong sense of prototypical emotions, the works are surprising, moving, delightful and interior experiential transitions are apparent. Thus it made sense to think about annotated responses in terms of evoked Valence and Arousal levels rather than discrete emotions.

For such works simple instruction protocols could be developed for participants, for example; (i) stand still and look around, (ii). Visit each piece in a specified order, (ii) sit in the centre of the space with eyes closed.

4.6.6 Experimental Setting |Example 6

Venue : The Lisson Gallery. London

Artist : Tatsuo Miyajima

Format : Electronic Sculpture.

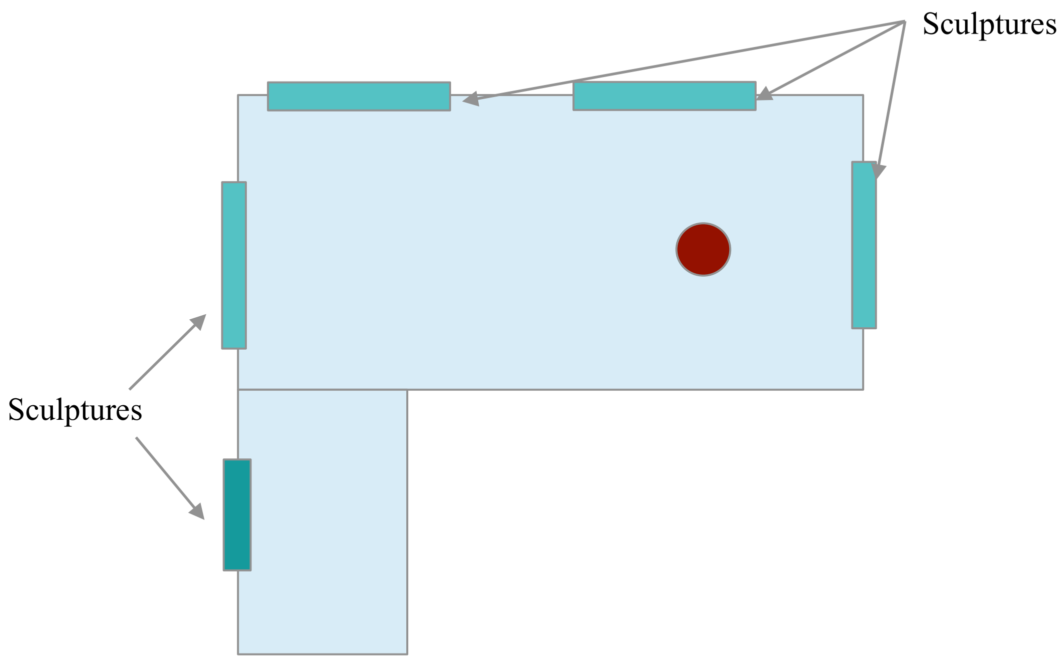


Figure 4.21: Topography of the gallery space for pilot study: Example 6. (red dot= optimal recording position)

On a separate occasion a test response recording was taken to Tatsuo Miyajima's solo show this time presented in the upper ground floor Lisson galleries (see Fig 4.21). In his works Miyajima uses the metaphor of cascading number series' comprised of variables 1-9, as representations of forms. In the exhibited works, with the aid of an Artificial Intelligence expert a series of panels with 'numerical communities' were presented. The works have a conceptual aesthetic, in that they provide a very cerebral process of analysis from which an emotional encounter may arise. However, it was very difficult to describe ones emotional responses to such a body of works, as subjectively one focuses on trying to interpret some form of meaning or associative comprehension. The emotional word that eventually arose was the term 'Cold'; this may have been due to the intellectualisation process, which leads to an emotional distance one feels to the work. After processing and classification, the recorded signal indicated higher levels of withdrawal or negative valence. This may in some way reflect the withdrawal process from the stimulus through forms of cerebral analysis rather than a negative experience. In such an instance the sensation of 'cold', can be a highly enjoyable experience.

Upon leaving the gallery a heightened emotional response was felt, which was very positive and arousing. It was considered that particular types of works might not present a strong immediate emotional response, but still produce a response that can be measured after a short delay. This again led to considerations of pre-post recordings, and also interval recordings & surveys that could be incorporated to gauge the emotional responses as the experience settles into memory over time.

Thus, as such a stimulus seemed to be affective in a different manner of emotional engagement to traditional emotional studies, it was felt that such a work, alongside a series of other works may operate best in forms of longitudinal studies where processes such as pre-post recordings, temporal reflections, alongside continuous recordings could be incorporated. In this way studies centred on predictive possibilities may also be explored.

4.6.7 Experimental Setting | Example 7

Venue: The Pace Gallery. London
Artist: James Turrell
Format: Light Installation.

A final test recording was made at the Pace Gallery, with another form of temporal stimulus. The Pace Gallery is a prestigious International Art Gallery that represents some of the worlds most highly established artists. Thus in this instance negotiations for requesting a recording were understandably lengthier than others and only permitted when the gallery was closed. This was to prevent any infringement on the experiences of members of the public and customers viewings.

James Turrel is an artist who uses 'light and indeterminate space to extend and enhance perception'. In this presentation an expansive gallery space was partitioned into three smaller rectangular spaces. At one end of each space was a cutaway recess in the wall. From a distance these looked like old television screens. Within these recesses a series of colour frequency and temperature transitions occur which perceptually distort the solidity and characteristics of the space, and at moments induce hallucinatory optical effects.

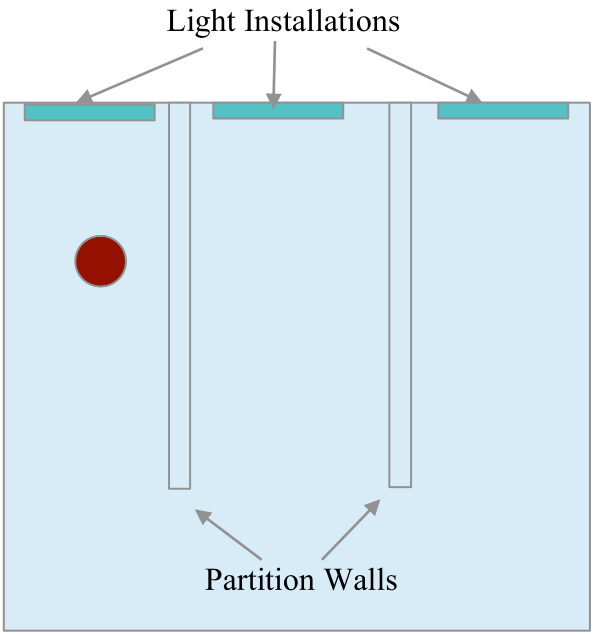


Figure 4.22 . Topography of the gallery space for pilot study: Example 7. (red dot= optimal recording position)

The work viewed was a newly constructed work Kermandec 2014. After starting the EEG and audio-recorder, the researcher explored the screen up close, and then backed away in stages to finally settle in front of the work at a distance of approx. 5 metres (see Fig 4.22). An effort was made to ensure that at each viewing location a static recording was made and also an oral annotation.

Subjectively, Kermandec has the ability to induce a real change in mood and emotion. It is increasingly relaxing, enthralling, and awe inspiring. One feels refreshed, quiet, invigorated and at times startled as apparitions begin to appear. It was felt that this would make a really good work for an experiment, especially due to the transition in pronounced feeling from the exterior public world to the impact of the work. Unfortunately the logistics and permission required to hold a group study here counter its potential.

4.7 Conclusion.

From this series of test recordings, an evaluation was made as to which form of cultural artefact and natural setting offered the most potential for conducting an experiment to meet the context and requirements of this foundational research project.

This consideration was in terms of (i) a real world setting, (ii) repeatability of stimulus and setting, (iii) a stimulus of sufficient strength to elicit strong and varied emotional responses, (iv) appropriate annotation method and (v) control of variables.

The range of works experienced highlighted a variety of complex rich aesthetic-emotional elicited experiences that were unique for each type of work, and also for each artist. Each artist communicated different sensibilities and modes of information transfer. Thus attempting to reduce these rich and complex experiences to simple information of valence and arousal levels through the single modality of EEG can in no way transparently fully describe these responses. Yet through our experiments by exploring one general facet of these experiences it may be possible to lay a foundation, upon which we may be able to iteratively explore with further scope and complexity.

Whilst it was felt that each offered a unique instance of investigation and insight into natural emotional responses, the Theatre setting was deemed most appropriate. It offered an excellent form of strong emotion elicitation, the natural minimization of a participant's movement, and a stimulus that was as exact for each participant as may be possible in a real world scenario. There was the factor of each performance having a unique audience who may respond with difference, yet again it was felt that these are reflective of a true naturalistic setting. Also whilst each participant may respond uniquely to the conveyed emotion by the performer, the presentation of stimulus was clear and intentional, and far less ambiguous than in the other works viewed. It was that felt for annotation purposes a post- experience survey and SAM test could be used to effectively label the recordings.

Regarding the potential of recording responses to Artworks in gallery spaces, it was felt that with group shows, such as the Light show at the Hayward, protocols of movements and timings could be inserted into the experiment, or if naturalistic randomness was required the participants left to wander freely. The annotation method of audio recording to gather immediate responses was highly effective.

For presentations of individual Artists works, it was felt that these might function best in a form of longitudinal study for single participants with considerations of working towards potential predicative systems. Here evaluations may be conducted for continuous, pre-post comparisons and temporally reflectivity.

Thus with sufficient consideration given to the context and nature of the work, effective strategies of investigation and experimental design can be constructed for gathering appropriate and meaningful data and annotation.

CHAPTER 5

Signal Processing

5.1 Introduction.

Primarily Brain Computer Interfaces (BCI) investigated assistive technological solutions for less-abled groups with particular emphases for sufferers of locked in syndrome (Cecotti 2010). The popularisation of this method of procuring a communication pathway between brainwaves and external devices has led to its wider adoption in applications beyond healthcare inclusive of marketing, creative fields, and also emotion research. There is a traditional BCI pipeline that is as follows;

(i)Signal Acquisition
(ii) Pre-processing.
(iii) Feature Extraction
(iv) Classification
(v) Application

Each context and project that uses this BCI pipeline configures each node of the methodology to suit its ends. In this chapter a transparency of the methods will be detailed and also the process by which such determinations were made.

5.2 Signal Acquisition.

For Signal Acquisition, we are using the Emotiv Epoch EEG headset. It has 14 bio-potential sensors with gold plated connectors, which conform to 10-20 International standard at sites AF3 (Fp1), AF4 (Fp2), F3, F4, FC5, FC6, F7, F8, T7, T8, P7, P8, O1, O2. It uses saline soaked felt pads that are placed in contact with the scalp. There are 2 electrode reference sensors for placement at both left and right mastoids. The Headset has a 2048 Hz internal sample rate, which is down sampled to 128 Hz. Its proprietary 2.4GHz wireless connection is via a custom USB receiver, and has specialized software Testbench for both viewing the live signal and recording live raw EEG data in the edf format. The Headset has a 2-axis gyroscope from which head movements can be detected. A Li-poly battery, 680 mAh, powers it and allows it to run continuously for 12 hours.

There are a number of different purchase options available, which include options for including their facial expression and affective state suites. The Education package we will be using for this project provides access to Raw EEG data.

5.3 Pre-Processing: Artefact Reduction Overview.

Within an EEG signal there is the potential for a variety of significant artefacts to manifest that can disguise its true readings. Artefacts may be introduced by the recorded subject, the recording device, or exterior electrical interferences. The most problematic of these in a still subject, are Ocular Artefacts (OA) that arise through blinking and eye movements. Endeavours have been made to realise processes that may reduce their potential distortion of the signal content. Following is a brief overview of the most popular methods used in artefact reduction (AR). This is followed by a practical testing of the most favourable methods, so that a process may be defined of use within this project.

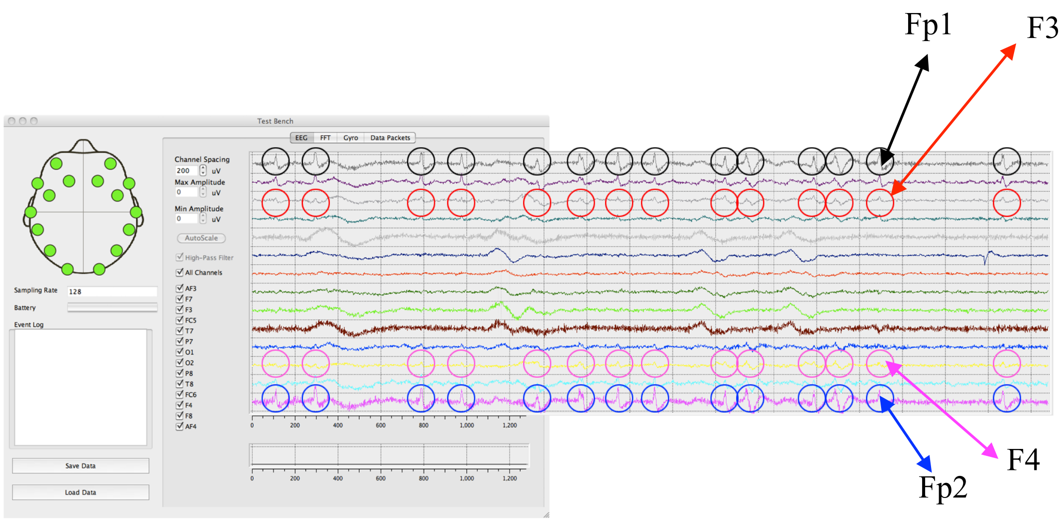


Figure 5.1: Illustration of the impact of blinking on a EEG signal recorded in the Testbench interface.

In figure (5.1) we can see an example of the residue that a eye-blink (artefact) leaves on a EEG Signal recorded in the emotiv Testbench Interface (circled). As is visible within this image, we can note that as the distance increases from the ocular region, the artefact becomes relatively reduced. Whilst the artefacts can be viewed as prominent in electrodes Fp1/Fp2 (above the brow) they are less distinct in electrodes F6/F7 (temple Lobe region), and further reduced in electrodes F3/F4 (frontal cortex region). As discussed else where in this study, for the context of this research, we are interested in detecting the emotional vectors of Valence and Arousal via electrode pairing F3/F4 in the Frontal lobe. So within this selection of electrodes, at the point of recording we have already managed to reduce some of the potential Ocular Artefacts (OA). However it is important to examine methods that have been developed and widely used, to see if any of these may be profitable for inclusion in this study.

Avoidance:

Avoidance is the process of attempting to eliminate artefacts generated by the subject at the initial point of the signal recording. This limiting is applied through instructing participants to remain still and not blink throughout the duration of the signal recording. Naturally, this is best implemented in very short recordings and in laboratory settings where such controls are easier to apply.

Rejection:

Rejection of Artefacts can occur in two ways. Firstly, any EEG Epoch (discrete recording) that can be identified as containing artefacts may be rejected outright from the study. The second approach is to identify and isolate artefact-ridden portions of the signal, and to extract only these from the signal. The signal minus artefacts may then be concatenated. Whilst this rejection technique is a solution, its discards valuable data and is not viable for studies using continuous recordings with a dynamic stimulus, as the timeframe is considerably altered.

Subtraction:

Subtraction is an artefact elimination method that uses further designated sensors to record a separate channel to capture the artefacts at their point of generation. This signal can then be used to subtract a weighted sum from each sensor relative to their location in the 10-20 international system.

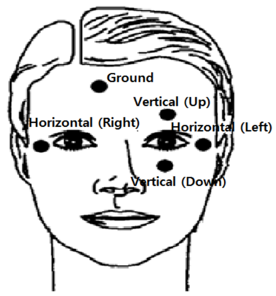


Figure 5.2: Traditional EOG electrode placements.

The Electro-oculogram (EOG) is a device/method which places sensors around the eyes at specific locations to record and measure the electrical potential/charge that is generated through eye movements and blinking (see figure 5.2). This technique may also be adapted for EMG, and ECG. As mentioned above, when their occurrence is noted, a subtraction assessment can be made for all electrodes. However this is not an effective technique for OA elimination in all contexts due to the overlap between brain data and OA signals in the EEG sensors placed over the Pre-frontal-Cortex (PFC) and Frontal-lobe (FL). Thus when weighted subtraction is applied, it inevitability extracts elements of the desired brain signal from the recording, changing the data at its first point.

Linear Filtering :

OA (Ocular Artefacts) < 4Hz
ECG (Heart Beat) == 1.2Hz (in the region of)
EMG (Muscles) > 30Hz
Electrical Interference > 50Hz

Figure 5.3: The Frequency ranges of physiological artefacts that may impact on the EEG signal.

If we take into consideration the known neural bandwidth ranges of the signal which the EEG detects 0 - upper 60 Hz and also consider the ranges that specific artefacts register in EEG signals (see Fig 5.3) we may assess that by filtering the signal within the range of 5-30 Hz we can potentially operate within a safe upper and lower oscillatory range of less prominent artefact registrations (Bos 2006, Mikhail,El-Ayat, Coan and Allen 2013).

It is important to note that the frequency ranges of the signal we are interested in for potential emotional detection are Alpha (8-13Hz) and of interest is Beta (13-30 Hz). Both lay in this 'Safe' region. Thus linear filtering holds great potential as a viable process for this research project.

Spatial Filtering :Independent Component Analysis (ICA)

Independent Component Analysis (ICA) is a form of Blind Source Separation (BSS) that attempts to locate independent signals within a mixed signal, to their points of generation. The most lucid explanation of this process is often presented as the cocktail party problem, where multiple microphones record a mixture of voices within a room, and the issue being to separate each voice. Applied to the context of Artefact Removal (AR), if all signals can be separated in this way, then there is the potential that one or more channels will hold all the artefacts that are desired for removal. Whilst this process calls for manual expert identifiers, specialised Matlab libraries such as EEGLAB, partially include this process within their interface. Yet again within this, the experimenter is required to manually select an ICA channel associated with OA for removal.

Thus whilst this technique has great potential for reducing artefacts, its major flaw is the manual identification of an artefact channel. Whilst in short time frames this maybe a justifiable pursuit, when we are dealing with lengthy or continuous signals this may not be viable. However, ICA has gained popularity over an earlier BSS algorithm Principle Component Analysis (PCA) as PCA assumes the extracted components to have orthogonal spatial topographies, where as ICA assumes statistical independence, which is more appropriate for EEG recordings (LeVan, Urrestarazu and Gotman, 2006).

5.3.1 Overview Summary

For this particular project that uses EEG to detect continuous emotions in naturalistic settings it is necessary to test applicable methods for their viability. Avoidance and rejection are not appropriate due to the real world setting within a continuous timeframe. Subtraction in terms of EoG is not desired as a process due to both its potential distraction in the experiment and its potential distortion of the signal (PFC & FL locations) at the first point of recording. However Linear Filtering and Independent Component Analysis both hold promise and as such it is important to test their viability on both a 'Test signal' and a 'real world' signal.

5.3.2 DC Offset Removal

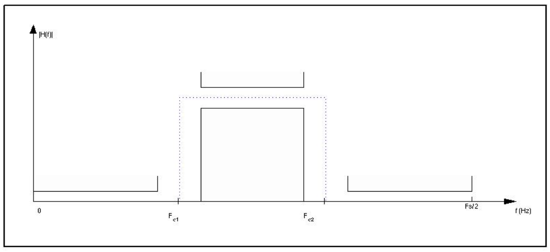


Figure 5.4: A Bandpass filter allows chosen spectral frequency values within a selected range to be retained whilst all others are discarded

The first step in this process is the removal of the dc offset inherent to the headset. This can be achieved through a high pass filter at 0.16 Hz, as stated in the manufactures headset instructions. This step can also be incorporated into a bandpass filter when extracting desired frequency ranges (see Fig 5.4). For this project Alpha (8-13 Hz) is the frequency range of interest, and Beta (13-30Hz) as a signal range of potential interest. Thus when required, the signals were passed through separate bandpass filter for each. To ensure a steep cut off slope and eliminate unwanted frequencies a 1000 order cut-off was used (see Fig 5.5).

5.3.3: Artefact Reduction Test 1: Linear Filtering

The researcher made a one-minute test EEG recording via the Testbench Interface. This was in natural conditions, sitting relaxed in front of a laptop, with no stimulus. This setting was at the researchers home that was noiseless and distraction free. During the one minute recording the researcher sat still and blinked intermittently. The researchers aim was to capture a signal recording where there may be a clear distinction between OA and non-OA parts of the recording.

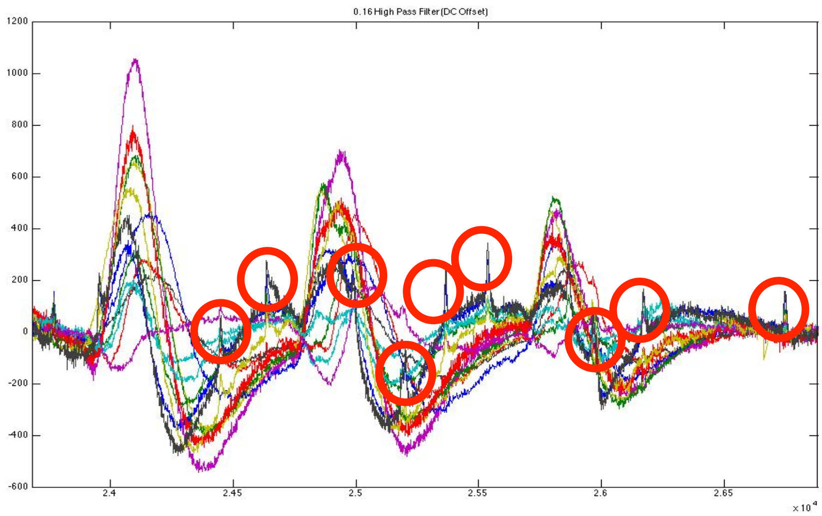
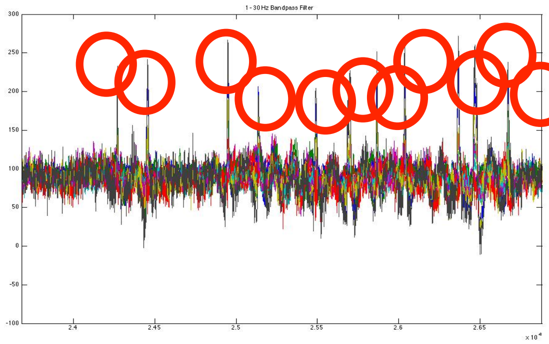
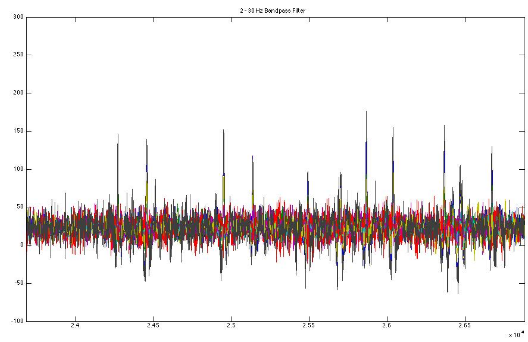


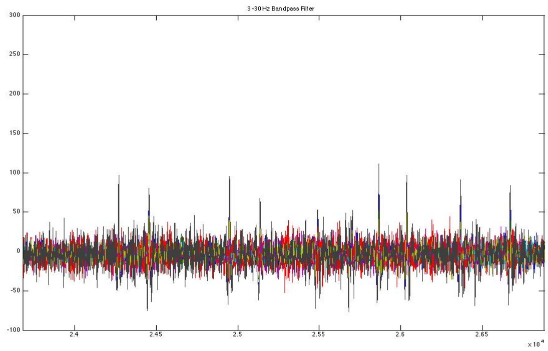
Figure 5.5: OA can be viewed in the EEG after a 0.16 Hz HighPass Filter (Dc Offset)

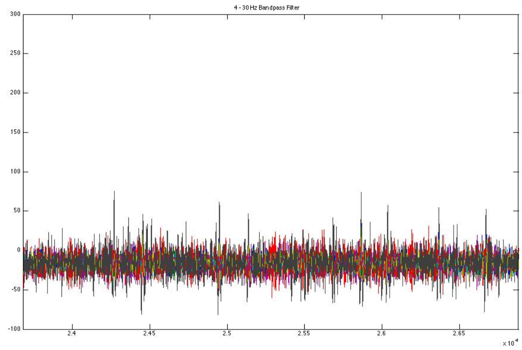
After filtering the DC offset, within the wild oscillation of the signal we are still able to visually discern and identify the majority of the Blink artefacts as circled in red (see Fig 5.5).

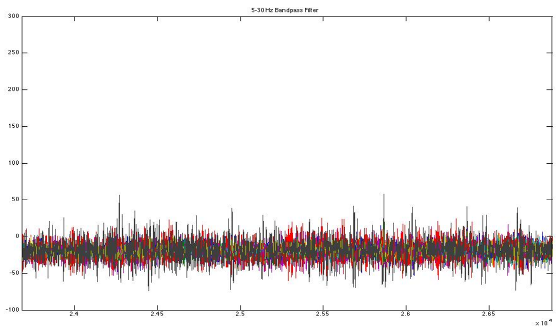
In order to test accounts that through filtering we may be able to remove many of the artefacts which operate at the lower frequency ranges, for example; OA < 4Hz, ECG == 1.2Hz , we can iteratively raise the lower floor level of our bandpass filter to discard specific frequency ranges.

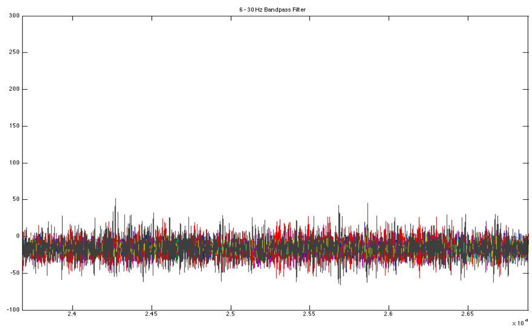

(1 - 30 Hz Bandpass filter)(Blinks circled)

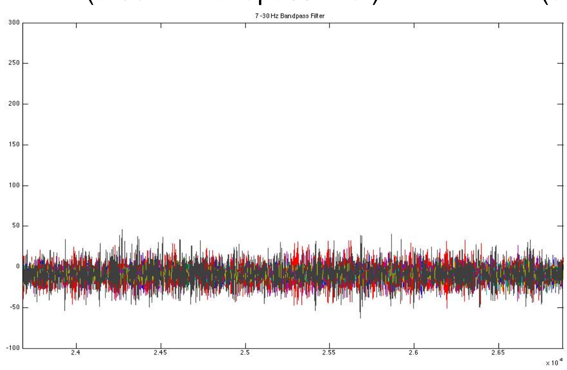

(2 -30 Hz Bandpass filter)


(3 -30 Hz Bandpass filter)


(4-30 Hz Bandpass filter)


(5-30 Hz Bandpass filter)


(6-30 Hz Bandpass filter)


(7-30 Hz Bandpass filter)

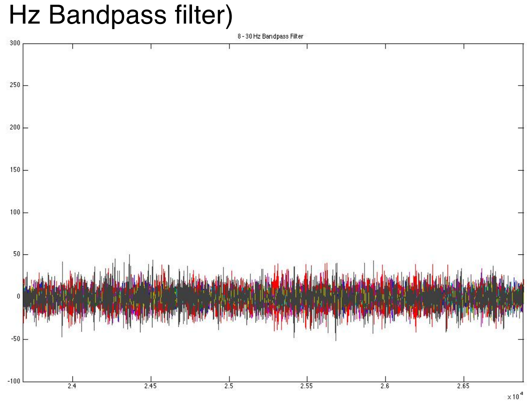
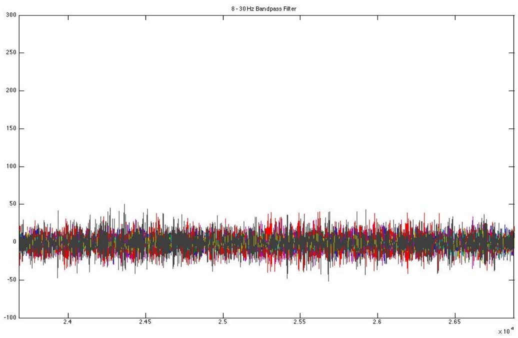

(8-30 Hz Bandpass filter)

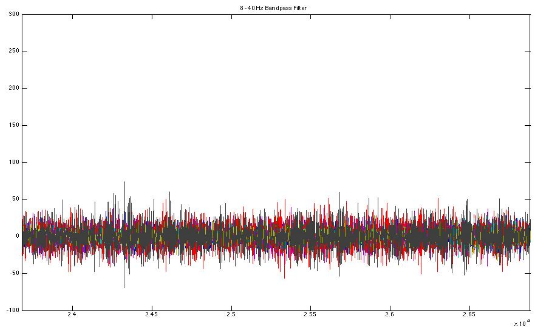
Figure 5.6: The OA become less prominent through linear filtering, as shown here by iterative raising the high-pass bandpass floor successively by 1 unit.

The preceding figure (see Fig 5.6) shows a series of bandpass filters for all electrodes. In each successive image the lower floor level of the bandpass filter is raised by 1Hz. In comparing the results between the 1-30Hz and 5-30Hz bandpass filters, we can see a vast reduction in the signal of the blink artefact. As a reminder it is important to note that the identified frequency ranges we are interested are Alpha, and of potential interest is Beta whose combined activity functions is the oscillatory region of 8-30 Hz.

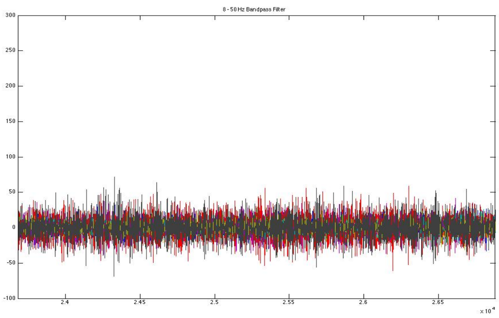
In the above sequence we have demonstrated artefact reduction through the heightening of the lower floor element of a bandpass filter, which as peer research proposes reduces OA (Bos, 2006, Mikhail et. al., 2013). We may also raise the ceiling of the bandpass filter beyond 30Hz, to see the extend of potential noisy artefacts of EMG (Muscles) > 30Hz) and Electrical Interference (> 50Hz) being introduced into our signal.



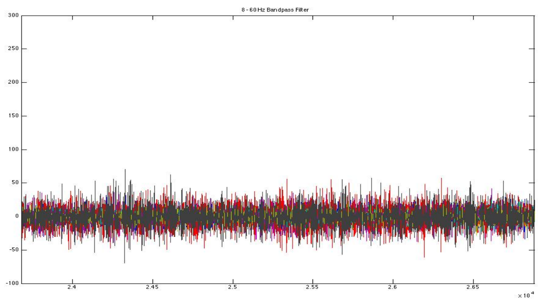
(8-30 Hz Bandpass Filter)



(8-40 Hz Bandpass Filter)



(8-50 Hz Bandpass Filter)



(8-50 Hz Bandpass Filter)

Figure 5.7: Iterative heightening of the bandpass ceiling by 10 units beyond 30 Hz does not show any dramatic changes in the EEG signal.

In figure 5.7, whilst on careful viewing we may be able to view differences, we cannot note a similar impact of raising the ceiling of the bandpass filter successively by

10 units, as we did with raising the floor value. If we attempt to raise this ceiling value beyond 63Hz we go beyond the measurable threshold of our acquired data signal.

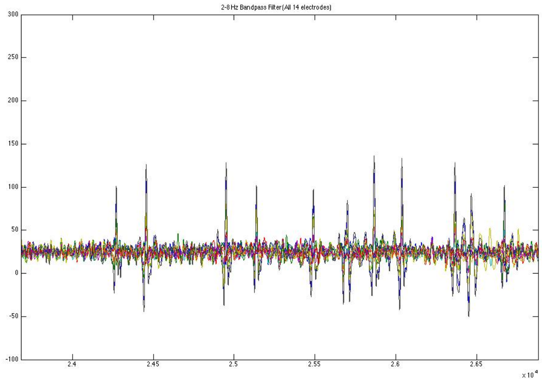


Figure 5.8: 2-8Hz Bandpass Filter (all electrodes)

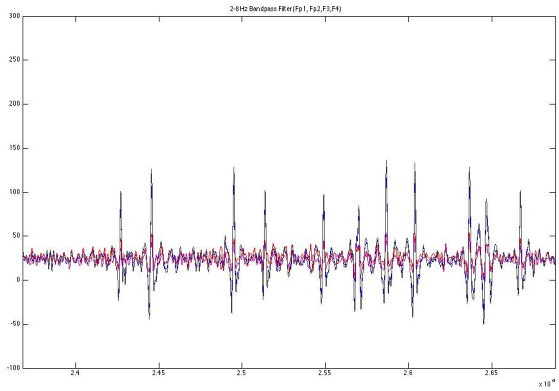


Figure 5.9 ; 2-8Hz Bandpass Filter (Fp1(blue),Fp2(black,F3(red),F4(pink))

In the above figures we can view a 2-8Hz bandpass filter for all electrodes (see Fig 5.8), and for only the electrodes in the regions of interest (see Fig 5.9), which confirm that the majority of the OA artefacts are contained within this oscillatory region of the signal.

Thus in the above test, we are able to see that by narrowing the bandwidth of our recorded signal to the bandwidths of interest (Alpha 8-13Hz, Beta 13-30Hz) we are managing to significantly reduce the more severe disturbances of our signals as reported in peer reviewed research. Whilst elements may still remain, we can potentially further reduce these through electrode selection, as Figure 5.9 shows, the artefacts are much more invasive to the signal in electrode pairing Fp1/Fp2 than in F3/F4.

5.3.4 Artefact Reduction Test 2: ICA test signal

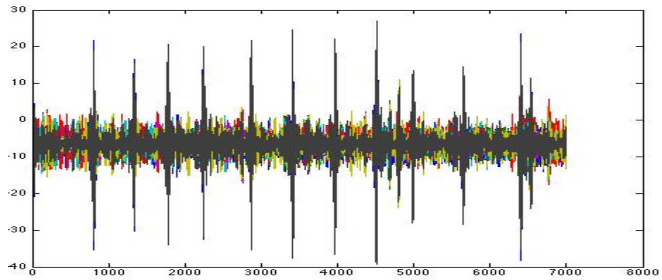


Figure 5.10: The blink test EEG signal. All electrodes (Bandpass filtered 8-13HZ)

For this test exploring the potential of Independent Component Analysis to further reduce OA, we are using the same signal used above in the linear filtering test (see 5.3.3). This has been bandpass filtered to extract the Alpha frequency (8-13Hz). Post filtering, we can note that there is still a residue of the blink artefact in our signal (see Fig 5.10). On closely inspection, despite eliminating a major factor of their impact in electrodes (Fp1/Fp2) there is still a residue albeit at a much lower rate. We can examine the extent of the artefact remaining in the signal through the visualisation of one hemispheres electrodes; Fp1, F7 and F3. (see Fig 5.11, below).

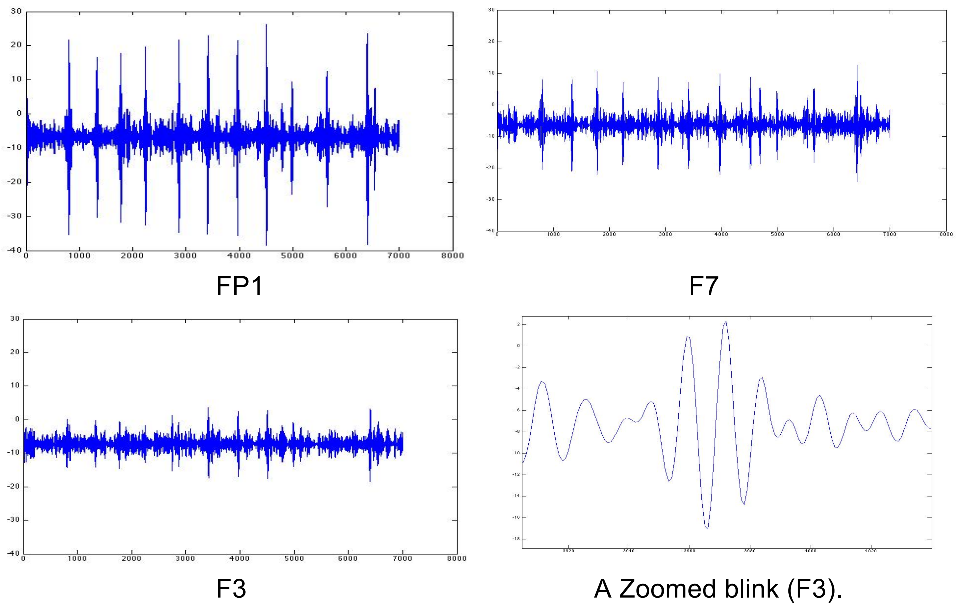


Figure 5.11: A closer view of post-linear filtering blink signal residue from electrodes, Fp1 (top left), F7 (top right), F3 (bottom left), and a close up of OA in electrode F3, (bottom right).

Figure 5.11, also allows us to confirm that the increased distance from the point of artefact generation reduces the artefacts interruption. In Fp1, the spikes are more significant than in F3 (paired with F4) which is the location we are interested in. In the bottom right corner of the image, we can see a detailed view of an OA from electrode F3. Here the blink can be seen to affect the signal for approximately 40 samples, which at a sample rate of 128 Hz translates as approximately 1/3 of a second.

Reviewing popular peer choices of ICA (Xiaowei et. al, 2010, Koelstra and Patras, 2013), the FastICA algorithm was selected and was available as a Matlab Library plugin for the apple mac OS. This was used in the following test to examine whether this smaller residue of OA could be further eliminated.

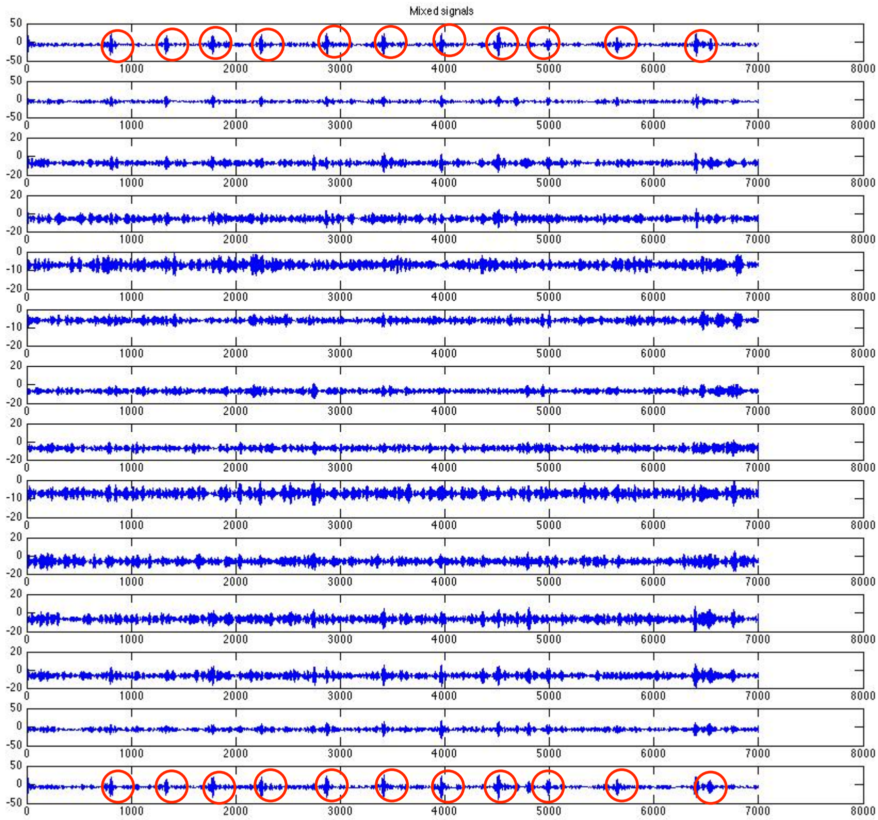


Figure 5.12: The original 14 electrode signals as viewed in the fastICA interface.

Firstly, within the FastICA interface we can plot the original signal to view the oscillations of each electrode channel. In figure 5.12, the blinks have been circled in red for visibility in channels Fp1/Fp2 (row 1,14). We can also visually detect their falloff rate in subsequent electrodes F6/F7 (row 2,13) and F3/F4 (row 3,12).

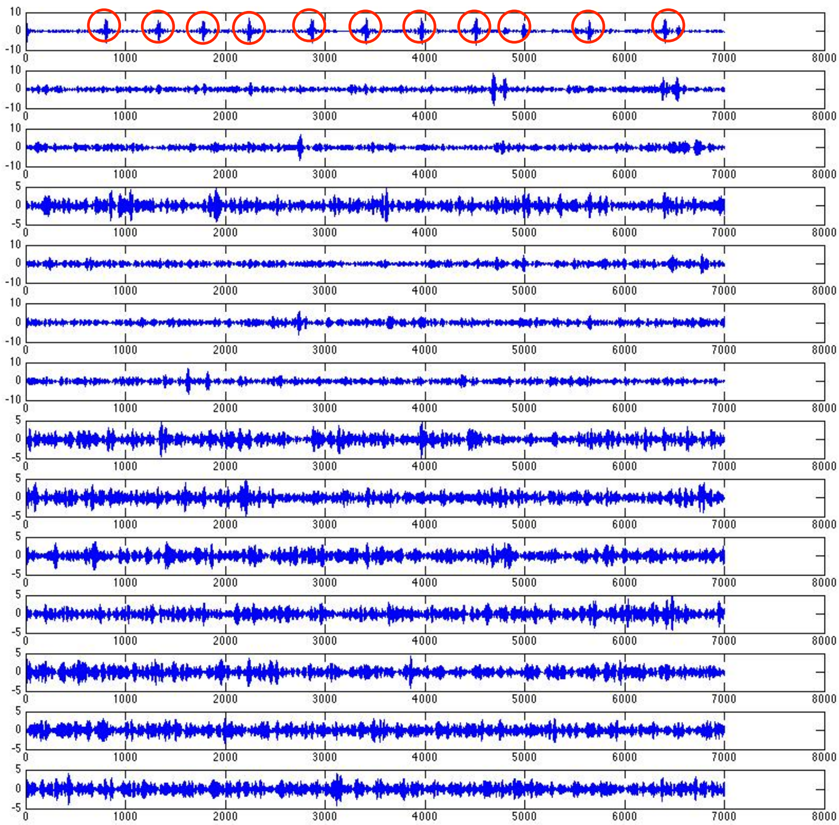
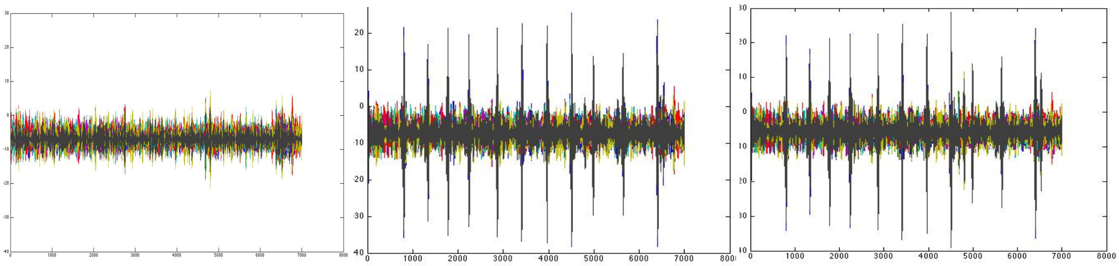


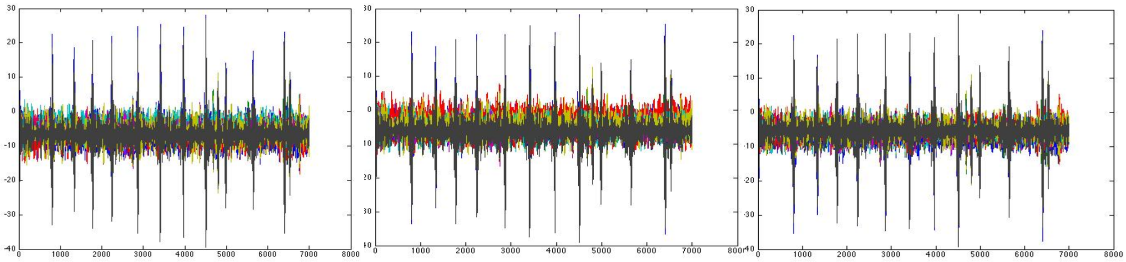
Figure 5.13: The calculated independent Components (by fastICA)

The FastICA algorithm was implemented on the whole signal. Figure 5.13, shows the independent components of the previously mixed signal after the FastICA algorithm had been applied. We can see that row 1 (blinks highlighted in red) contains the component signal that corresponds to the blink artefacts. However making such a decision with this scale of image may be problematic for even the more experienced identifier. Thus Matlab's plotting tools can be enlisted to assist in this identification, by providing a process that allows a closer inspection to ensure we locate the correct component of the signal to extract.

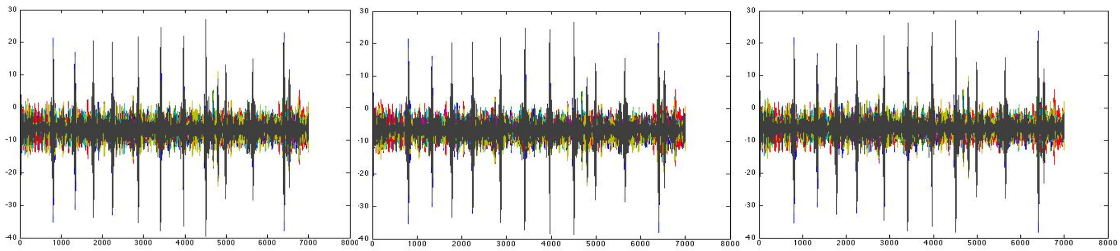
Using the Independent Component data generated by the algorithm, we can create 14 instances of the separated component matrix (we have 14 components due to the headset having 14 electrodes.). For each successive instance of the matrix, we can zero out one component, before remixing the individual component matrices to output 14 different post-ICA signals. Here in each remix matrix we have eliminated one component. Thus we are able to produce 14 post-ICA plots that we can now visually scale for closer viewing inspection. (see Fig 5.14, below)



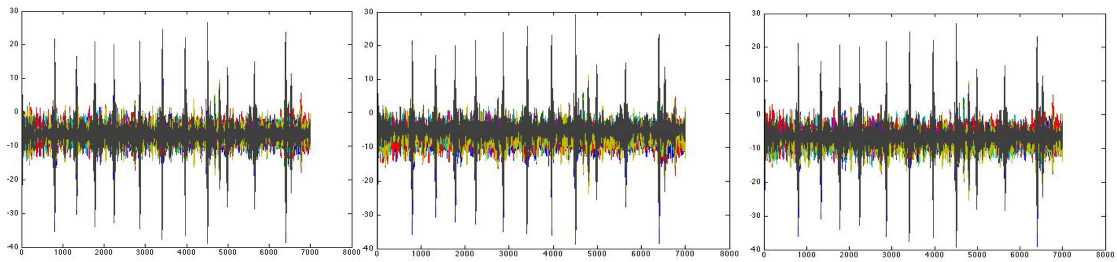
Remixed Matrices with components removed (-1,-2,-3) (l-r)



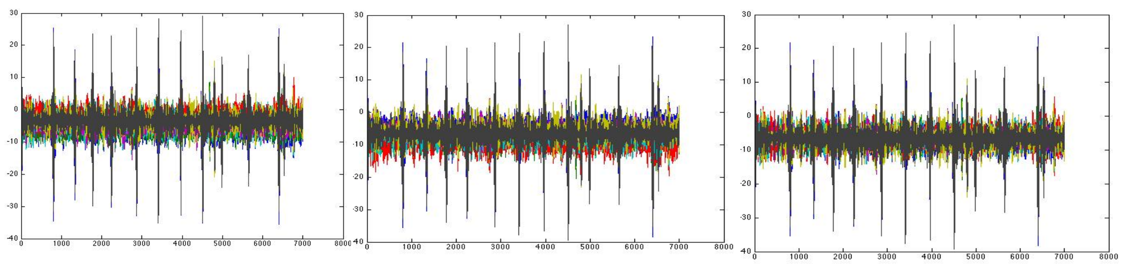
Remixed Matrices with components removed (-4,-5,-6)(l-r)



Remixed Matrices with components removed (-7,-8,-9(l-r)



Remixed Matrices with components removed (-10,-11,-12)(l-r)



Remixed Matrices with components removed (-13,-14),Original signal (pre-ICA)(l-r)

Figure 5.14: This figure show the 14 different remix matrices, each with one component removed. In this way we can clearly determine which ICA component holds the OA.

In consideration of figure 5.14 we can clearly detect that of the 14 components the first image (top row, top left) is the independent component which holds the blink signals, and corresponds to the visual image as displayed within the FastICA interface (see Fig 5.13).

Using this visualisation technique we can examine the impact of removing the undesirable component has on our signal in further detail, and also for any electrode we feel is necessary. However, whilst we may agree that the FastICA algorithm can potentially remove artefacts from the signal, it is also important to ensure that we are not inadvertently compromising the integrity of our data signal.

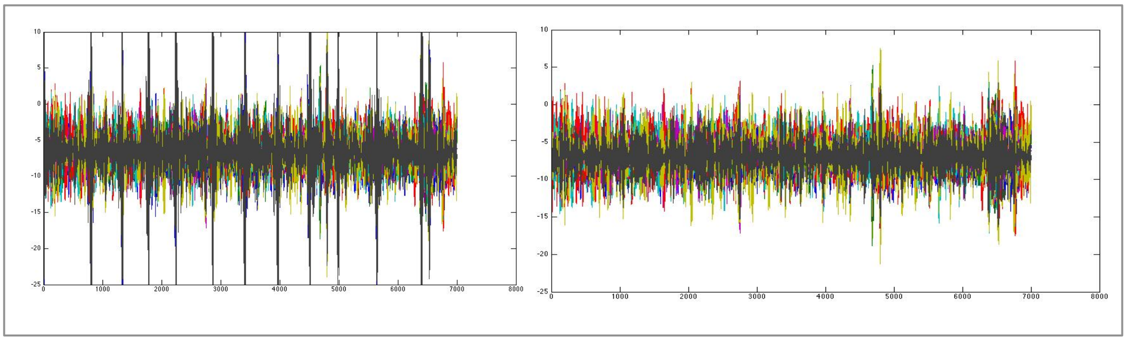


Figure 5.15: (left) pre-ICA, (right) post-ICA signals for all electrodes.

Figure 5.15, above, shows the pre-ICA signal & the post-ICA signal for the collective 14 electrodes. Whilst we can note that the OA's seem to have disappeared, we cannot detect whether the signal integrity has remained intact.

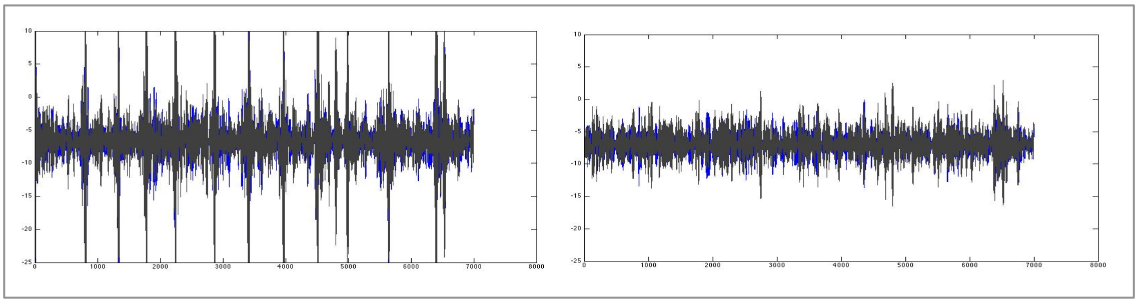


Figure 5.16: (left) pre-ICA, (right) post-ICA signals for electrodes Fp1 & Fp2.

In the above image (see Fig 5.16) we can view a dramatic reduction in the OA spikes relating to the electrodes at locations Fp1 & Fp2.

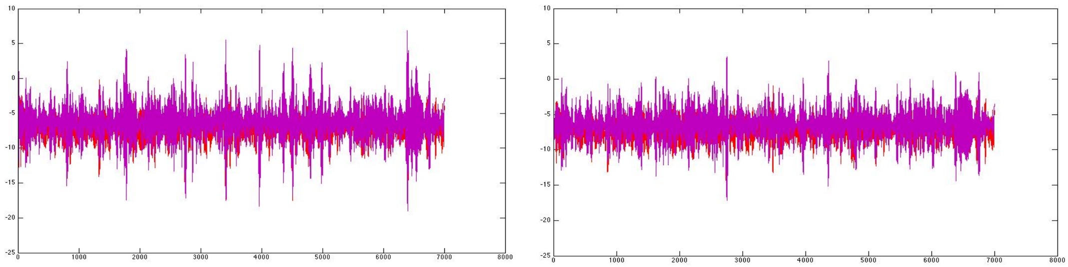


Figure 5.17: (left) pre-ICA, (right) post-ICA signals for electrodes F3 & F4.

In figure 5.17, we may note, a lowering of the OA spikes in electrodes located at position F3/F4. As this is our electrode location of interest, we can look even closer at a smaller portion of our signal.

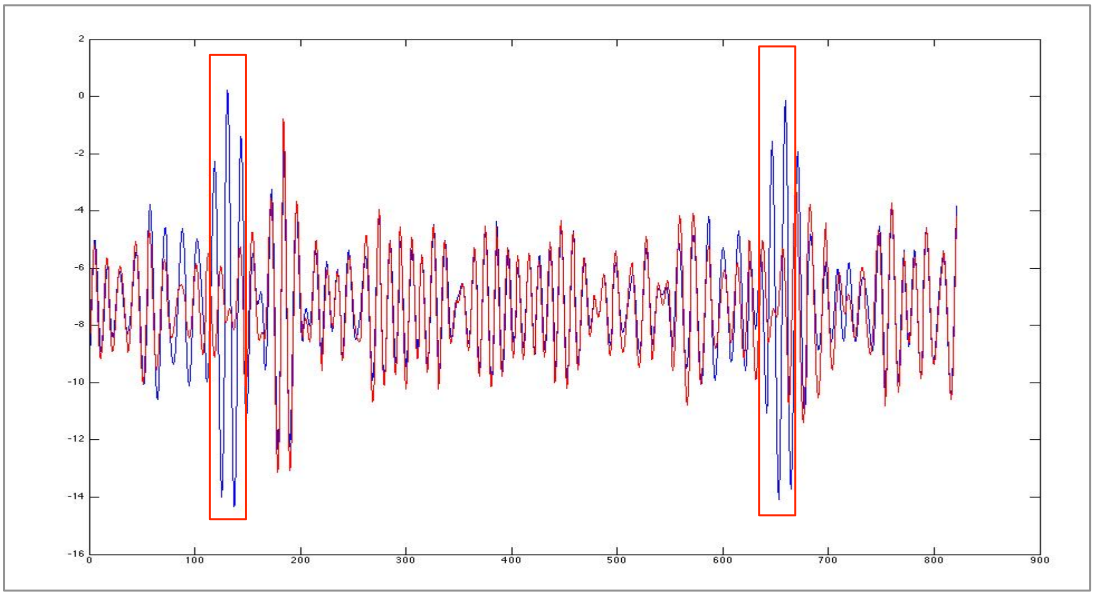


Figure 5.18; pre-ICA (blue) and post-ICA (red) signals for electrode F3 (detail).

In the above figure (see Fig 5.18) we can clearly see that the signal is reduced at the points where OA are clearly situated (marked by red rectangle). However we can also note that there is also a change in the signal at all points, which serves to change our data. This can be made more explicit by looking at a OA free section in more detail.

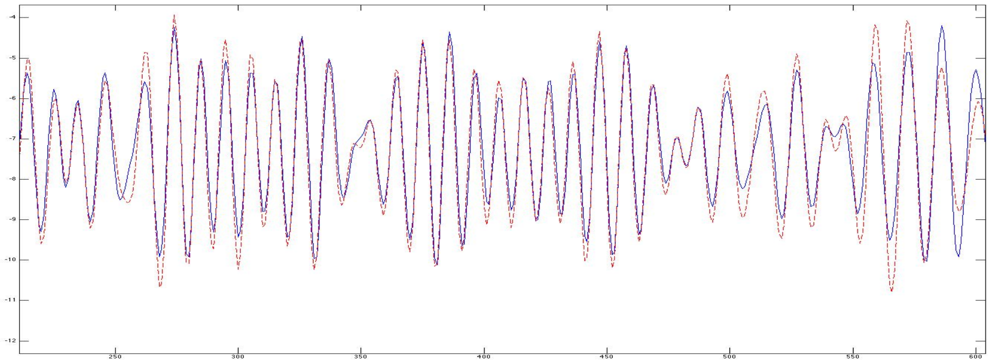


Figure 5.19 : OA free section detail for electrode F3, pre-ICA (blue), post-ICA (red)

Figure 5.19, shows the impact of the FastICA algorithm on a OA free section of our signal. Whilst the impact seems nominal we can clearly see that at each point, the signal has changed. Thus we are losing the signal's integrity, and abstracting from our original signal. Naturally with any such removal process there maybe a playoff between removing artefacts in a signal and retaining the original signal. Thus in making a decision whether it is beneficial to use ICA we can statistically examine the changes in the signal for the electrodes we are interested in.

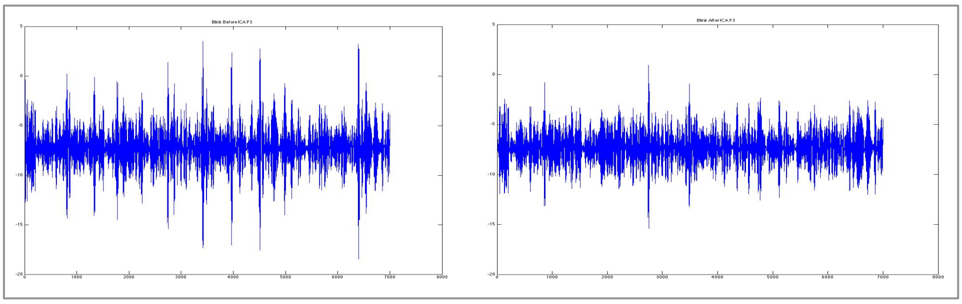


Figure 5.20 : pre-ICA(left), F3 post-ICA (right) for electrode F3.

Electrode F3	Statistic	Before ICA	After ICA	Difference
	Min	-18.47	-15.38	-3.09
	Max	3.5	0.92	2.57
	Mean	-7.25	-7.33	0.07
	Std	2.06	1.65	0.53
	Variance	4.23	2.71	1.53
	Sum	-5.07E+004	-5.13E+004	563

Table 5.1: The Statistical summary for electrode F3, pre/post ICA.

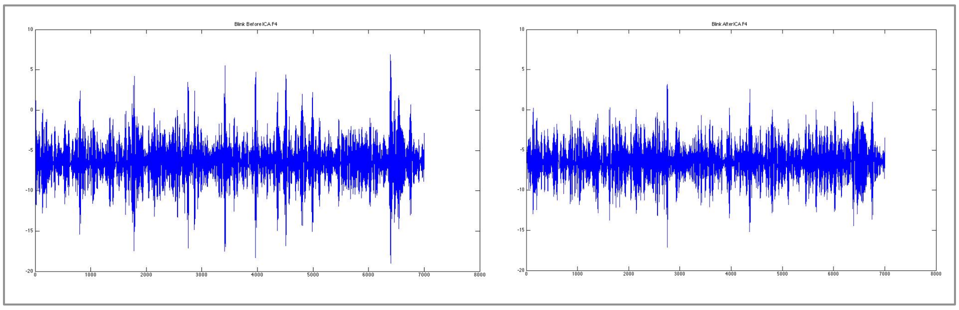


Figure 5.21: F4 pre-ICA(left), F4 post-ICA (right).

Electrode F4	Statistic	Before ICA	After ICA	Difference
	Min	-19.02	-17.14	-1.88
	Max	6.88	3.17	3.7
	Mean	-6.4	-6.5	0.09
	Std	2.6	2.23	0.38
	Var	6.78	4.96	1.82
	Sum	-4.4840e+04	-4.55E+004	651

Table 5.2: The Statistical summary for electrode F4, pre/post ICA.

In the above tables 5.1, and 5.2, and in figures 5.20 and 5.21, we can note that there is a difference in all values. We can note that for the mean of the signal (which is the measure we will be using in further analysis throughout this project), there is a difference of 0.0748 for electrode F3 before and after ICA, and a difference of 0.0929 in the mean value before and after ICA for electrode F4.

We may select an OA free portion of the signal and inspect it numerically to assess the impact that ICA has on our recorded signal.

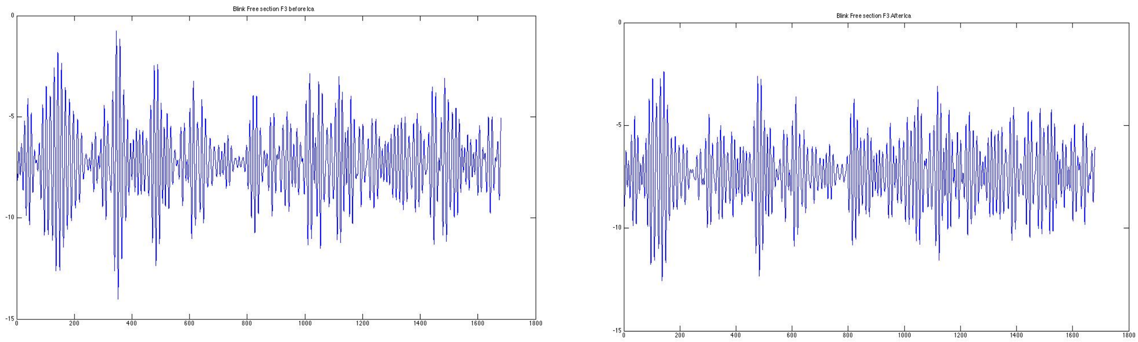


Figure 5.22: This figure shows a section of the signal, which is known to be OA free, for electrode F3 pre-ICA (left), post-ICA (right)

Electrode F3 (OA free)	Statistic	Before ICA	After ICA	Difference
	Min	-14.03	-12.56	-1.46
	Max	-0.76	-2.37	1.61
	Mean	-7.25	-7.34	0.09
	Std	1.72	1.51	0.21
	Var	2.95	2.28	0.67
	Sum	-1.2187e+04	-1.23E+004	148

Table 5.3: The Statistical summary for electrode F3, pre/post ICA (OA free region).

Electrode F4 (OA free)	Statistic	Before ICA	After ICA	Difference
	Min	-15.05	-12.82	-2.24
	Max	2.23	-0.02	-2.24
	Mean	-6.4	-6.49	0.1
	Std	2.31	2.17	0.14
	Var	5.34	2.17	0.63
	Sum	-1.0753e+04	-1.09E+004	161

Table 5.4: The Statistical summary for electrode F4 pre/post ICA (OA free region).

We can gather the mean values obtained in tables 5.1, 5.2, 5.3 and 5.4, to clearly examine the extent of the impact of FastICA algorithm on our signal

Electrode F3	mean difference value prior and post ICA
whole signal including OA	0.07
OA free signal portion	0.09

Table 5.5: The mean value comparison for electrode F3, which reveal the extent of change to the whole signal and the OA free region.

Electrode F4	mean difference value prior and post ICA
whole signal including OA	0.09
OA free signal portion	0.1

Table 5.6: The mean value comparison for electrode F4, which reveal the extent of change to the whole signal and the OA free region.

Tables 5.5 and 5.6 confirm that we are changing the integrity of our signal, especially in the regions of the signal that contain no OA's.

5.3.5 Artefact Reduction Test 3: ICA Real world signal.

The above ICA process test was performed on a test signal, and thus it is important to also test this on a signal that parallels the experimental set up to be used in this research project. A 'real data' signal was captured in the chosen experimental set up of a live theatre performance audience. The experimenter replaced the participant for this recording. The chosen venue was the Soho Theatre, Soho. London, and the performance viewed was Bitch Boxer by Chloe Jackson (see chapter 4).

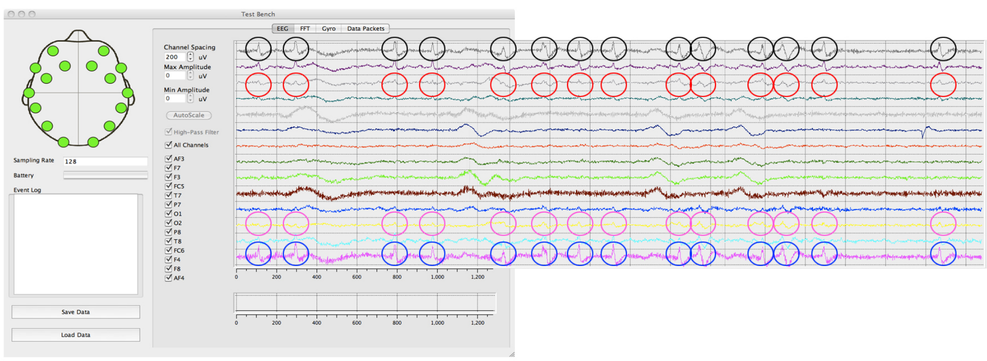


Figure 5.23: This figure shows a section of the signal (as described above), that contains a series of OA's (circled in black/Fp1, red/F3, magenta/F4, and blue/Fp2)

The Original signal was recorded in emotiv's specialized EEG recording interface testbench. The signal was replayed within this interface and the timings of clustered and single OA's noted. The signal was segmented in to artefact and non-artefact regions by these notes. For this analysis a section (length 3200 sample, 25 seconds) containing OA was chosen

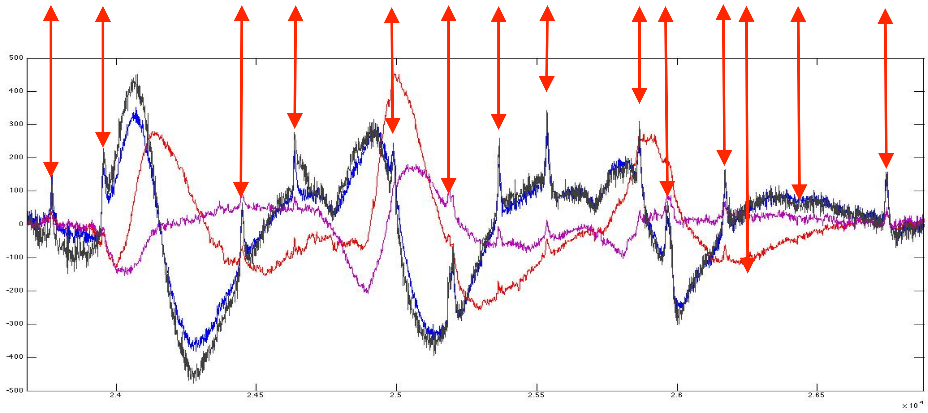


Figure 5.24: The extracted signal with dc offset (0.16) removal.(black/Fp1, red/F3, magenta/F4, and blue/Fp2) (blink position are highlighted by the red arrows.)

The same procedure as outlined for the test signal was used. Further the visualisation of the ICA data was also enlisted for easier viewing and selection in Matlab as outlined above (see section 5.3.4).

Electrode F3	**Statistic**	**Before ICA**	**After ICA**	**Difference**
	Min	-20.55	-18.8197	-1.7323
	Max	4.2615	2.9808	1.2807
	Mean	-7.2806	-7.1312	-0.1494
	Std	3.0088	2.9666	0.0422
	Var	9.0526	8.8010	0.2516
	Sum	-2.3298e+04	-2.2820e+04	-478

Table 5.7: The statistical summary for electrode F3, pre/post ICA.

Electrode F4	**Statistic**	**Before ICA**	**After ICA**	**Difference**
	Min	-16.82	-15.29	-1.53
	Max	3.82	3.18	0.64
	Mean	-6.33	-6	-0.33
	Std	2.86	2.64	0.22
	Var	8.2	6.98	1.23
	Sum	-2.03E+004	-1.92E+004	-1056

Table 5.8: The statistical summary for electrode F4, pre/post ICA.

In the returned results as with our test signal, we noted that artefacts were heavily reduced in electrodes pairing Fp1/Fp2, and importantly that we were also losing and degrading the signal in our electrode pairing of focus, F3/F4 (see tables 5.7 & 5.8). Whilst for Fp1/Fp2 the mean value for our signal changed by a factor of -0.7805/-1.4146 respectively, in our electrode signal of interest F3/F4 it has changed by -0.1494/-0.3299.

Electrode F3 (0A FREE)	Statistic	Before ICA	After ICA	Difference
	Min	-11.3906	-13.0613	-0.9087
	Max	0.0090	0.6232	0.6142
	Mean	-7.1484	-6.4304	-0.7180
	Std	2.6051	2.6122	-0.0071
	Var	6.784	6.8237	-0.0373
	Sum	-2.5091e+03	-2.2571e+03	-252

Table 5.9: The Statistical summary for electrode F3 pre/post ICA (OA free section).

Electrode F4 (0A FREE)	Statistic	Before ICA	After ICA	Difference
	Min	-11.3906	-11.6809	0.2903
	Max	-1.6713	-0.2686	-1.4027
	Mean	-6.3910	-5.7721	-0.6189
	Std	2.1216	2.2151	-0.0935
	Var	4.5012	4.9066	-0.4054
	Sum	-2.2433e+03	-2.0260e+03	-217.3

Table 5.10: The Statistical summary for electrode F4 pre/post ICA (OA free section).

Tables 5.9 and 5.10, show the same test performed on a known artefact free region whose length comprised of 351 samples for electrodes F3/F4. As with the test sample we can also see that in the real world signal we are changing our signal at all points. For F3 the difference of the mean value for this section before and after ICA is -0.7180, whilst for electrode F4 the difference is -0.6189. These differences for electrodes F3/F4 are gathered below (see tables 5.11 and 5.12).

Electrode F3	mean difference value prior and post ICA
whole signal including OA	-0.1494
OA free signal portion	-0.7180

Table 5.11: The mean value comparison for electrode F3, which reveal the extent of change to the whole signal and the OA free region.

Electrode F4	mean difference value prior and post ICA
whole signal including OA	-0.33
OA free signal portion	-0.6189

Table 5.12: The mean value comparison for electrode F4, which reveal the extent of change to the whole signal and the OA free region.

This same procedure was also tested for the extracted Beta (13-30) signal. When attempting to follow the same process with the same signal portion restricted to Beta oscillations (13-30 Hz) we found that the lower cut off of the bandpass filter, shows that the OA were not outstanding within the signal. As determined within the above filtering test (see section 5.3.3) we can denote that many of the OA appear to operate at lower frequencies, which we determined by examining the whole signal and the electrodes of interest (Fp1,Fp2, and in particular F3,F4).

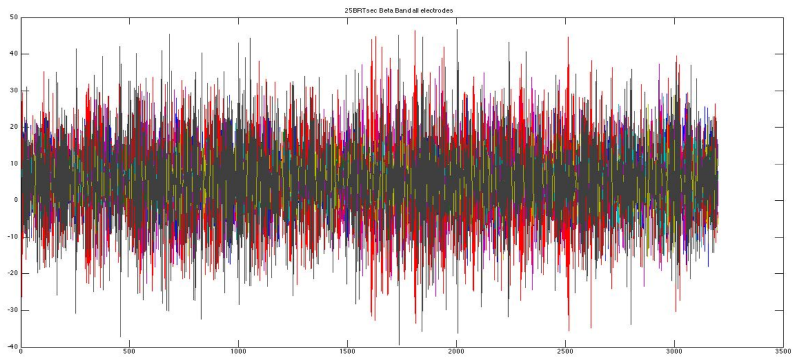


Figure 5.25: This figure shows that OA's are not visible in the test segment for Beta (13-30 Hz).

We found that it is not possible to clearly delineate OA as separate from any region of the signal for all electrodes (see figure 5.25). We can look in turn at the electrodes of interest that are closest to the ocular region that may register more clear prominence.

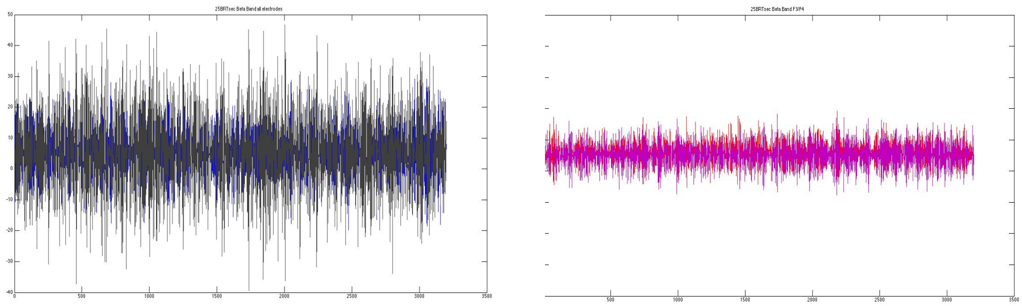


Figure 5.26: The sample Beta signal (13-30 Hz) , for electrodes Fp1/Fp2 pairing (left), F3/F4 pairing (right)

In figure 5.26, we find that for electrode pairings Fp1/Fp2 and F3/F4 at the beta band frequency, we are still unable to differentiate OA's from any region of the signal

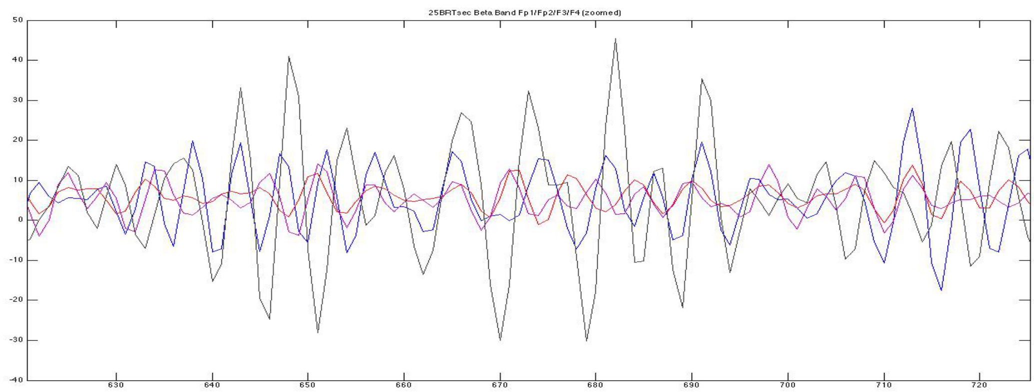


Figure 5.27: This figures show electrodes Fp1 (black), Fp2(Blue), F3(red), F4(pink) for the Beta frequency(13-30 Hz) for a region of the signal known to contain OA.

If we closely examine a region of the signal where a known OA is present we may view any OA impact on the beta frequency band of the signal. Further we may also note whether the signals from each paring of electrodes affect the other. As we can see from the above image (see figure 5.27) there is independence between them. The higher spikes of Fp1/Fp2 positions do not lead the F3/F4 paired positions signals, thus concluding that by using electrode pairing F3/F4 in the beta band range we do not have any major contamination of the signal by OA's. Whilst Beta was included in the above considerations as a frequency of interest, moving forwards we concentrated solely on the Alpha frequency due to the support we found in our literature review (see Chapter 3).

5.3.6 Potential Solution: Time Stamping.

Thus whilst the FastICA algorithm may be effective in reducing OA's, if it is used on a whole signal then there is a risk of the signal being transformed which may lead to miss-classifications. One option is to find the precise location and duration of OA's and to export them from the signal for ICA OA component extraction, before re-concatenating the signal. Kumar, Arumuganathan, Sivakumar and Vimal (2008), propose such a solution enlisting Discrete Wavelet Transform (DWT) to plot the timing of blink, which can then be reduced by "adaptive thresholding". Whilst to implement an arbitrary threshold may not be of value in this instance, we can briefly examine any potential in this method.

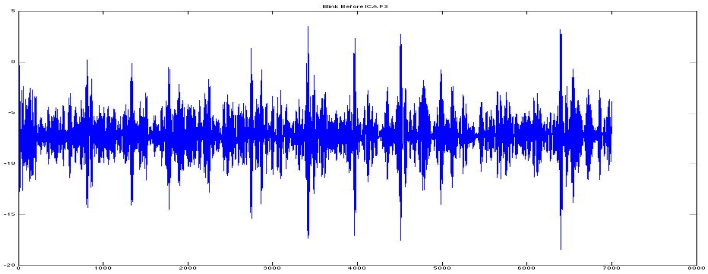


Figure 5.28: The blink test signal (+ 1000, preceding samples) for electrode F3.

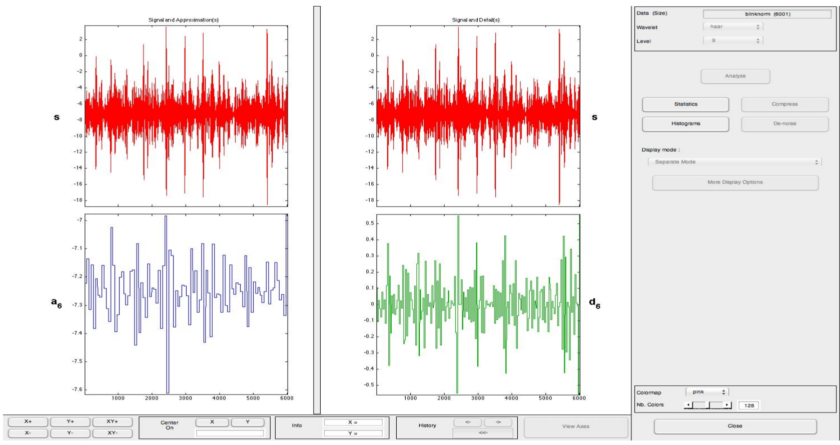


Figure 5.29: A screenshot of the DWT processing at level 6 approximation and detail, for electrode F3.

In principle DWT is a process of down sampling a signal, In figure 5.28, we can view the signal used in this test, which is the signal used in the above blink test (see section 5.3.3), with 1000 preceding samples included. Figure 5.29, shows the signal at the 6th level of approximation and detail (DWT). Here we can view the potential of being able to timestamp OA's through their breech of a given threshold in comparison to OA free regions. By exporting this time stamping to our initial signal, we may then precisely extract each OA occurrence within our signal, for individual ICA implementation. However we then have to manually perform the fastICA extraction process (as described above section 5.3.4), for each OA. When we consider the practicality of this as a process for a noisy signal in continuous timeframes of up to 1 hour, we may surmise that it is not affective or economical for this particular project. As we have determined, through filtering and electrode selection we have reduced OA’s to an acceptable level for the requirements of this research.

It should however be noted that proposals for automating ICA are an area of interest for the EEG community and as such, Viola et. al. (2009) present CORRMAP an semi automated IC identifier plugin for EEGLAB, which clusters ICA through thresholding. Winkler, Haufe and Tangermann (2011) envisage linear classification techniques to automate the selection of the ICA components to be extracted based on multiple statistical features. A similar approach is the usage of Bayesian Classifiers to separate decomposed ICA EEG epochs by LeVan et. al (2006) whereby signals are determined as EEG as opposed to Artefact through their registration under certain thresholds, which also accounts for both EoG and EMG.

5.3.7 Artefact Reduction: Conclusion and Discussion.

Firstly it is important to state our intentions. We desire to obtain EEG signals in natural settings from which we aim to decipher emotional responses to cultural artefacts. This is in order to gain some insight into these subjective emotional experiences. The EEG signal is sensitive to distortion from a variety of artefacts introduced by either the participant, the recording device or exterior electrical interference.

Above we have considered potential peer tactics for their removal, and tested those most suitable for this projects context, namely linear filtering, Independent Component Analysis, and also electrode selection.

Firstly above we have seen: through linear filtering of the signal to the oscillatory range we are interested in Alpha (8-13 Hz) and also of curiosity Beta (13-

30Hz) we are able to reduce the main body of artefacts of OA, ECG, EMG and electrical interference. Through electrode selection of F3/F4 we are further reducing any residual registration of OA's on our signal. Further we have explored the potential of performing the FastICA algorithm on our signal to further remove any impact of OA. We have found that unless we perform this algorithm solely on the OA region, we are compromising the integrity of our signal. Whilst a possible solution has been demonstrated using DWT, to timestamp their occurrences within our signal, this would still result in the manual performance of selecting the OA component for each selection.

For this project this is not deemed as efficient or affective as a process, as it is envisages that these signals will be lengthy and look towards real-time solutions. Further it is felt by the researcher that through the above mentioned process any potential artefacts will have already been reduced to an acceptable nominal level. Thus moving forward Linear Filtering to Alpha 8-13 Hz, and electrode pairing F3/F4, will be the most effective and efficient forms of reducing the presence of artefacts in our signal.

5.4 DFT: Window Length

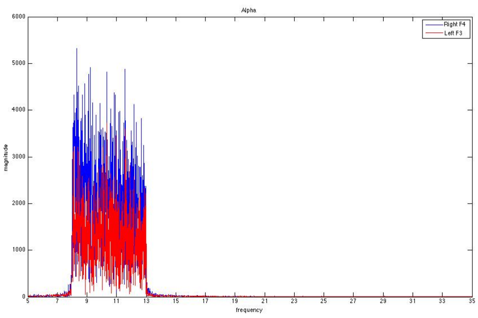


Figure: 5.30: Alpha FFT 8-13 Hz

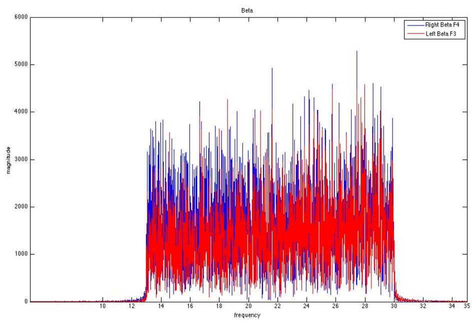


Figure 5.31: Beta FFT 13-30Hz

As previously mentioned, for this project Alpha (8-13 Hz) is the frequency range of interest, with an interest for Beta (13-30Hz). Thus as required, the signals were passed through separate bandpass filters for each. To ensure a steep cut off slope and eliminate unwanted frequencies a 1000 order cut-off was used. The temporal data signals can be made meaningful for analysis by translating them into the spectral frequency domain through two types of Fourier transform. A Fast Fourier Transform (FFT) can be applied to the whole signal to return intensity values for each of its frequency steps (see figs 5.30 and 5.31).

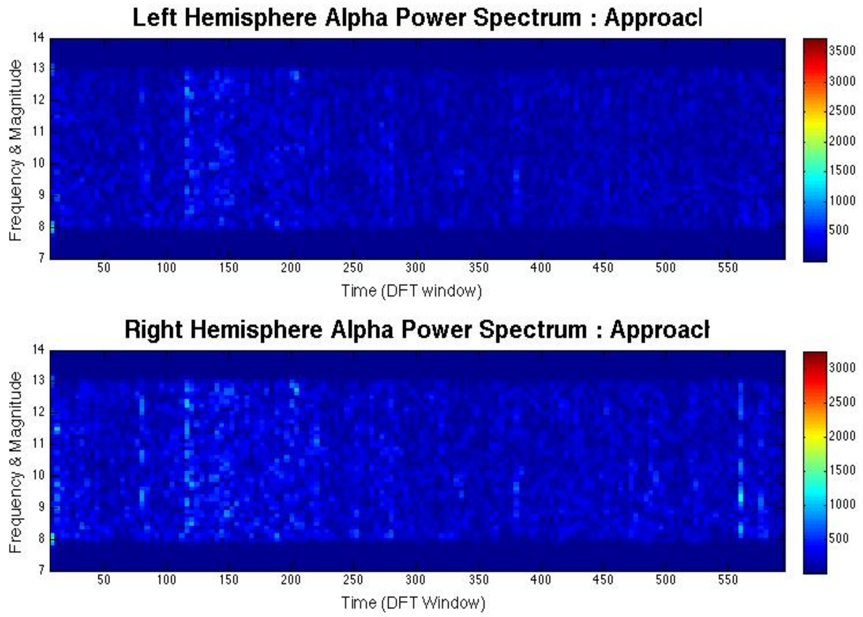


Fig 5.32: A spectrograph allows the spectral power to be viewed in discrete time units.

Secondly using a Discrete Fourier Transform (DFT), we can segment the signal into discrete window bins of any chosen length. The DFT window length selected for this project is a ; hanning window of 1024 samples (8 seconds) with 50% overlap to produce a spectral frequency representation of the signal for every sequential 4 second segment. In figure 5.32 we can see an example of a resulting spectrograph as plotted in MATLAB 2012b, This displays both time (X axis) and, frequency (Y axis), and magnitude (colour).

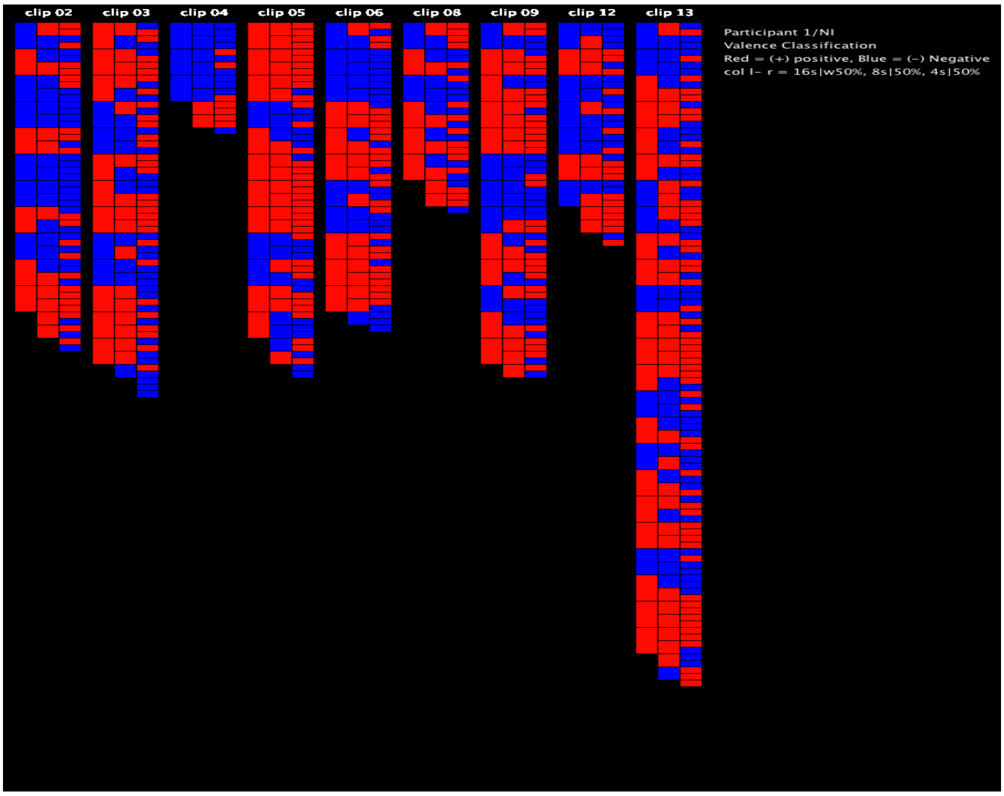


Figure 5.33. A visualisation of binary classification for Valence shows the correlations between 3 different DFT window lengths for nine separate stimuli responses for a single participant. Here each column represents the Valence response for a single stimulus. Within this 3 further columns represent the different window lengths of (left to right): 8 seconds, (16s|50%|128Hz), 4 seconds (8s|50%|128Hz) and 2 seconds (4s|50%|128Hz).

Tests were conducted into the selection of a meaningful window length. In figure 5.33 (below) we may view a representation of a resulting Valence classification; Red for (+), Blue for (-) , for participant (NI) in experiment 1 (see chapter 6). Each of the nine columns represents a stream of data for a single stimulus clip trail. Each column is further segmented verticality to represent the 3 window lengths tested; (left to right) 8 seconds (16s|50%|128Hz), 4 seconds (8s|50%|128Hz) and 2 seconds (4s|50%|128Hz). Horizontally we can see their temporal correlation for the trails duration. Visually, we may be able to detect slight differences in their classifications due to the averaging of different quantities of data, yet the majority of the class patterns integrity is kept intact. Naturally larger windows give a more generalised interpretation, whilst smaller window lengths present a more detailed view, accounting for these slight differences. It was felt that a 4 second value was appropriate in that it gave a good level of detail, and also a timescale that may prove valuable and meaningful in noting any changes in emotional dimensions

5.5 Data Labelling & Classification

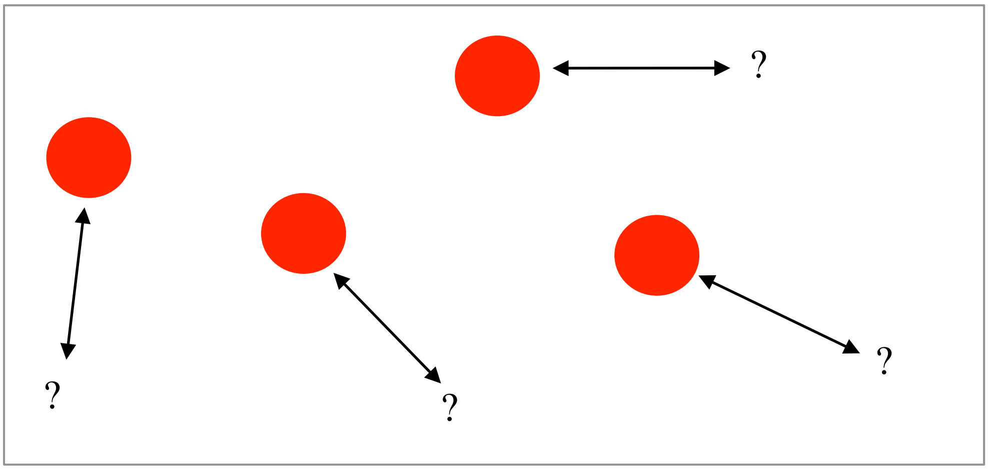


Figure 5.34. The Accurate labelling of data is vital for classification purposes.

5.5.1 Introduction.

As pointed out by Healey (2011) accurate data labelling is the biggest challenge facing affective computing, in the understanding of a data signal, and in turn then potentially being able to understand further incoming data. As per design of the experiment 1 (see chapter 6, and Appendix A; A.4), for part of the study survey, participants were asked to give a short annotative written description of any changes in their felt emotions over the course of the sound clip. This was to test whether meaningful temporal relationships may be observed in the analysis between the signal and their written annotations.

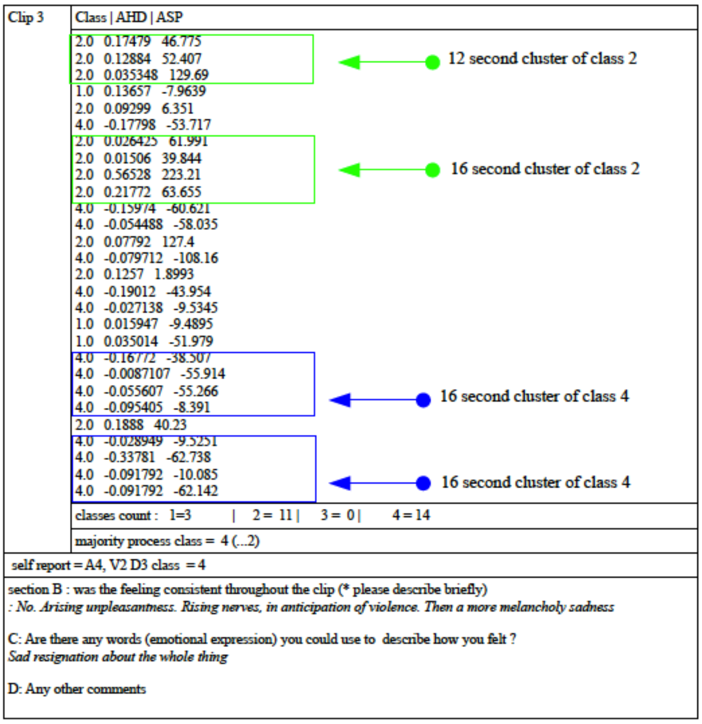


Figure 5.35: An illustration of speculative temporal correlations between participants written annotations and their stream of signal classification, which were computed every 4 seconds.

In the above example (see figure 5.35) we can infer a movement in the signal's temporal classification results which occur every 4 seconds, that matches the felt emotional movement suggested in the annotation. The participant annotated their experience as an inconsistent feeling, *" Arising unpleasantness. Rising in nerves in anticipation of violence. Then a more melancholy sadness".* At the beginning of the signal classification results, clusters amounting to 28 seconds for a sustained positive Valence, with low Arousal were noted. Towards the end of the signal, this transitions to a sustained 32 second clustering of negative Valence with high Arousal levels. Whilst this may infer a relationship between annotation and signal, it is important to consider that the annotation may not be of sufficient detail to have confidence in this example. This form of analysis considered these forms of relationships over the whole groups responses between signals and surveys. Whilst the number of correlations noted by the researcher exceeds a random occurrence, this form of analysis does not present the form of formal objectivity required. Whilst this may be overcome by having groups of independent identifiers note their own correlations, it was felt that this process would need further consideration, to be implemented as a justified objective classification method.

Thus, various classification methods were explored. Algorithms such as Support Vector Machine (SVM), K-Nearest Neighbour (K-NN) were considered and tested for suitability using Fishers iris data, looking at 'multi-class', in a one versus all operation and also 'binary' classification. When considering these as potential classification methods for this research projects data, the issue arose of a non-stationary temporal data signal against proposed singular ground truth values, where no complete relational confidence could be stated. Further for the two types of settings and stimulus used in the research, a definitive ground truth may not be procured, which again does not provide the required confidence.

Thus rule base EEG classification methods such as Baseline Correction classification were also explored for potential appropriateness in this project

5.5.2 Baseline Correction.

Baseline correction is a procedure that is ubiquitously used in EEG-ERP experiments. The principle of the process is to extract all elements of a signal that are not considered to solely be the response to an experiential stimulus, so that the only the response remains as the measure. This is achieved through recording a pre-stimulus baseline, whose mean value is then extracted from the signal, leaving only the measurable stimulus response (Woodman, 2010). Alternatively a baseline may be constructed from the mean of all related data signals, and then subtracted from each data point in the time series. Again this leaves only the response signal in relation to the baseline condition.

Whilst ERP studies are conducted in minute time frames, we can take the principle of baseline correction and apply it to our study in order to configure a suitable classification procedure that may be used within the context of this project.

Recording a pre-stimulus EEG baseline and subtracting its mean value from every data point in our experiment signal, we are left with a temporal signal of data points. Each data points then represents a dynamic positive or negative movement away from a zero point (the baseline mean). Further, it may also be possible to consider the distance from zero at each point as a signifier of the strength of the response for any given data point. By conducting this process for all participants in a study, it may also be possible to clearly plot and gauge their comparative responses at any given time across the participant group.

Thus through this method we may firstly arrive at sequential positive/ negative values for each vector of Valence and Arousal or a single averaged value for each epoch. Secondly we may composite their sequential or averaged values onto a 2 dimensional circumplex model, whereby they can be assigned class labels.

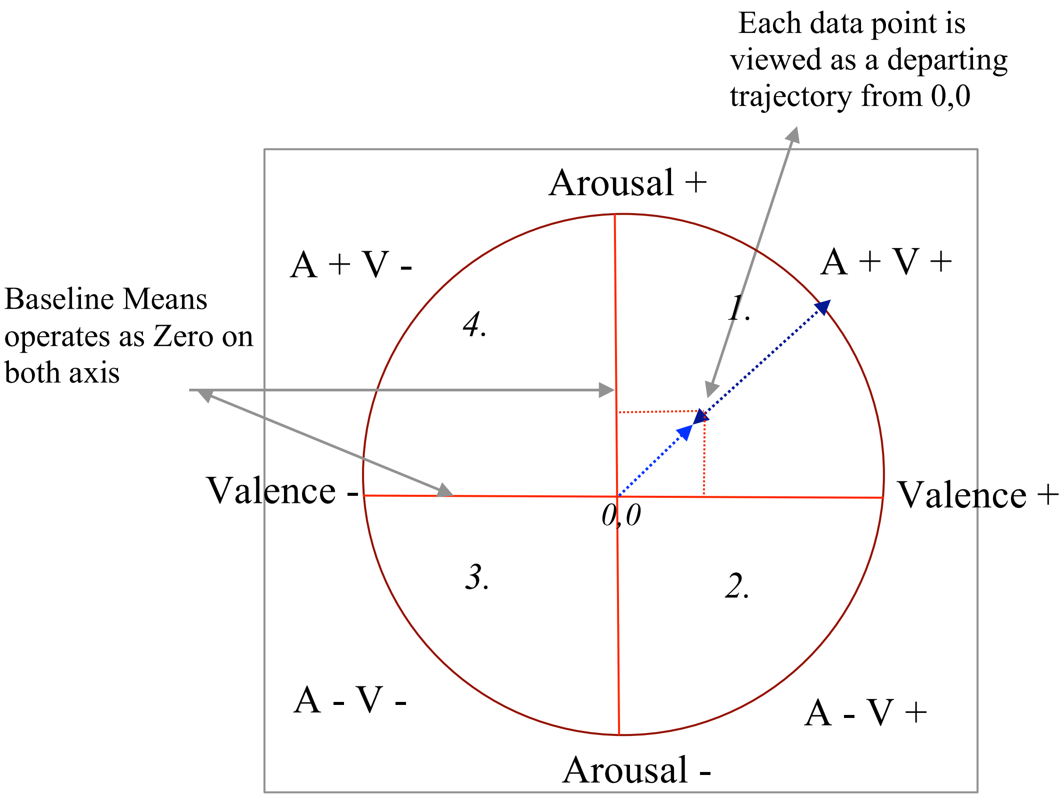


Figure 5.36. An illustration of the classification space and model.

Figure 5.36, presents the visualisation of this method (space). We can see that the baseline mean operates as a zero axis on both vectors of Valence and Arousal, which serve to categorise a two- dimensional space into four categories.

4 V- A+	1 V+ A+
3 V- A-	2 V+ A-

Table 5.13: The correlation table between Valence and Arousal values and Class Labels.

Thus each data point can be labelled as residing in one the four sectors dependent on the its composite Valence and Arousal value, as detailed in the table 5.13.

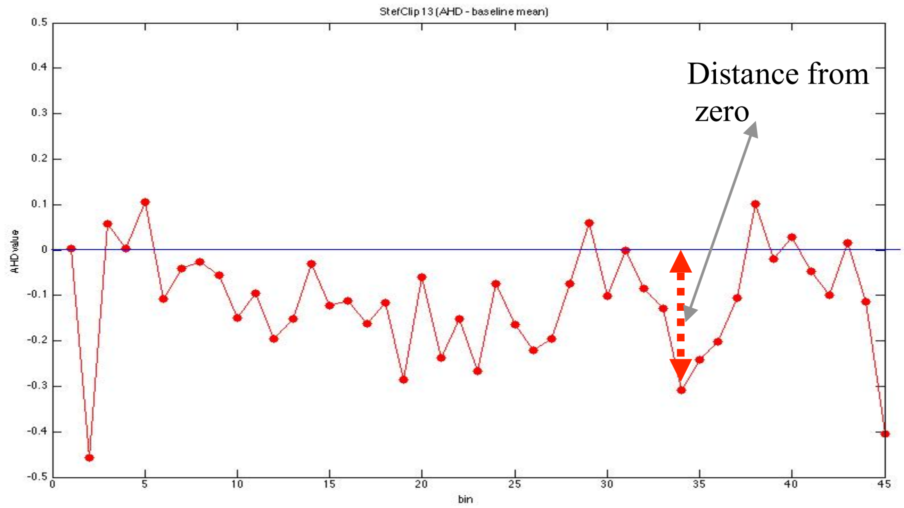


Figure 5.37: The sequential data points after the baseline mean has been subtracted for the Valence vector

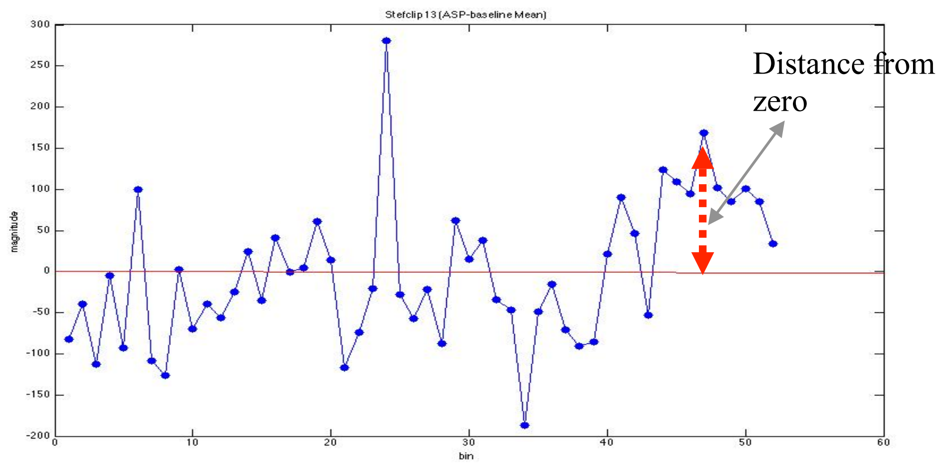


Figure 5.38: The sequential data points after the baseline mean has been subtracted for the Arousal vector

Figures 5.37 & 5.38, clarify this process for each vector, to show the results of baseline correction on a temporal signal. For this purpose, a participant's data from the experiment detailed in chapter 6 (see Chapter 6), has been used. Figure 5.37 represents the Valence vector, which is constructed through the AHD algorithm. Thus values less than zero represent negative Valence and a negative movement from the baseline, whilst greater than zero values represent positive Valence. For Arousal this is inversed where movements to lower regions indicate increased excitation, and higher regions represent increased relaxation (see Fig 5.38). As the above figures demonstrate, we can view the location of each temporal data point as a response, against the zero axes.

5.5.3 Cross Participant Charting

Within this method there is the potential for temporal cross-participant evaluation. Above in figures 5.37 & 5.38, the temporal vectors are shown for a single participant. By subtracting each participant's particular baseline from their particular signal in the same manner, it becomes possible to chart responses across a group of participants. In this way it becomes possible to find correlative responses at any particular given time, which potentially in terms of evaluating responses to a cultural artefact, may be of value

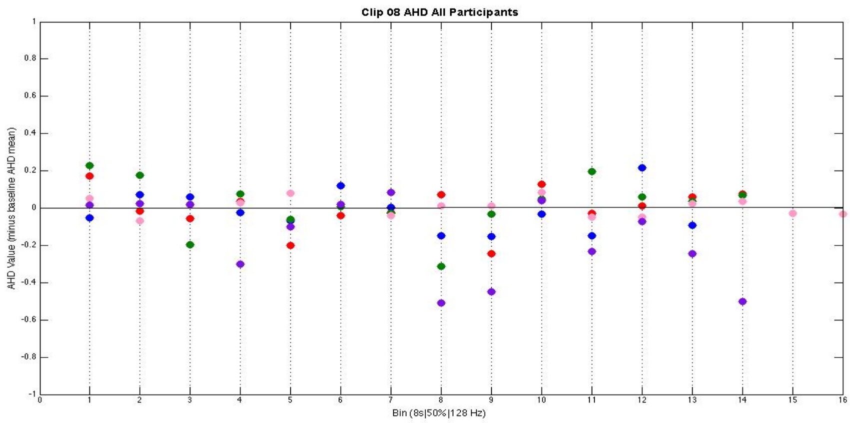


Figure 5.39: Multiple participants plotting for the Valence vector.

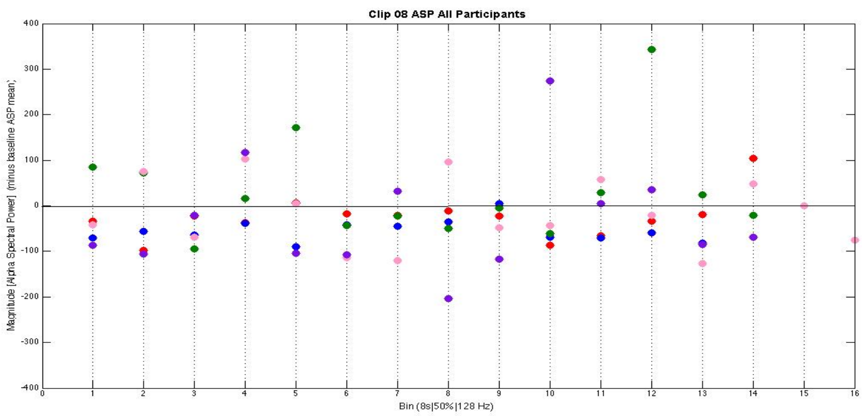


Figure 5.40: Multiple participants plotting for the Arousal vector.

CLIP 8	Valence SAM	signal + registration	signal - registration	dominant category	Arousal SAM	signal + registration	signal - registration	dominant category
P1(SN)	3(-/+)	(+)7	(-)7	(-/+)	3(-/+)	(+)2	(-)12	(+)
P2(SM)	5(+)	(+)5	(-)8	(-)	4 (+)	(+)1	(-)12	(+)
P3(SA)	4(+)	(+)9	(5)	(+)	2 (-)	(+)7	(-)7	(-/+)
P4(ST)	4(+)	(+)10	(-)6	(+)	4 (+)	(+)7	(-)9	(+)
P5(SO)	5(+)	(+)6	(-)8	(-)	4 (+)	(+)5	(-)9	(+)

Table 5.14: Shows the correlations between SAM rating, and the dominant signal registration category for Valence and Arousal over time.

Figures 5.39 and 5.40 alongside table 5.14 demonstrate how through this process, we may be able to both visually and numerically extract data patterns from our signal across participants at any given time Further we may also potentially classify data streams as being skewed towards either positive or negative dominance by summing up the data points registrations for each binary condition on either vector.

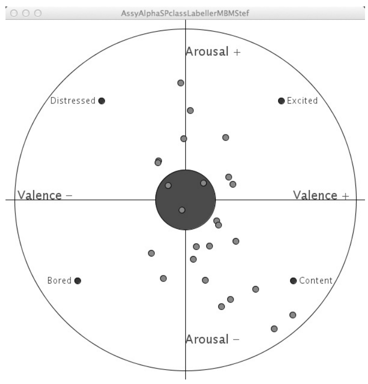


Figure 5.41: The classification space, created in the Processing environment

Through the above process the individual vectors, either as absolute values or reduced signifiers, could then be combined or composited as (x) and (y) values for classification within a 4 category, 2 dimensional circumplex space dependent on their location.

For test purposes the researcher chose to export the vectors to the Processing-programming environment where they were automatically given class labels, plotted in a visual representation (see figure 5.41) and made available for creative data visualisation purposes.

Whilst this process hold much promise for creative and artistic purposes, in this particular project, we are trying to validate a formal process, to compare the signal versus annotations. Thus as stated above, we do not have the level of annotation description available to attempt to validate the process, and this issue of temporal annotation requires further consideration. Thus as an concrete process, we may take a mean value for any designated region, or response clip as a whole, and compare this to a single SAM test variable. In both cases signal and survey may be reduced to binary signifiers, of either positive (+) or negative (-) value.

5.6 Conclusions.

Above we have attempted to give a transparent rendition of the processes by which we arrived at our methodological selection of signal processing; signal acquisition, pre-processing, feature extraction and classification, which will be used in our formal experiments.

For signal acquisition, we will be using the emotiv epoch headset, as this enables mobile raw EEG recordings to be taken. For pre-processing and AR removal, we explored a range of processes and found the best option for this project to be linear filtering. This was the most efficient and allowed for us to be able to extract the frequencies in the region of interest (Alpha, F3/F4), and simultaneously to eliminate the majority of artefacts without compromising or abstracting our signal.

For feature extraction we will be using AHD for valence, and ASP for arousal (see chapter 3). A single continuous feature vector will be extracted for each dimension. From these, the particular regions relating to each stimulus will be isolated.

For classification, there is the issue of comparing labels to signals. Regarding annotation labels, the subject's self- reports will produce a single SAM test value for each vector of Valence ad Arousal for each stimulus. Whilst it would be highly desirable to have a more detailed temporal annotation method, this becomes problematic in terms of accuracy. As the participants may apply these labels after the event, asking for detailed annotations might present a higher probability of mislabelling.

As stated above, using the Baseline Correction method we can extract only the stimulus response data that we require, and then average the EEG epoch relating to the each stimulus. This may be considered as an absolute value, or reduced to either a positive or negative signifier for binary classification. These averaged values will be compared to the SAM values for each stimulus.

It may be visible that throughout this thesis we have continuously seen the gap widen between our original propositions of gauging emotional responses to artworks, and what is attainable both in terms of method and technology. Naturally we have reconfigured our scope within these found limitations.

The methodological processes outlined above may be seen as part of a foundational testing to discern whether we may obtain on par results between laboratory and natural settings. If this is the case, then it may be possible to return to developing these methods into a more detailed temporal format. This would allow us to explore some of the different forms of potential aesthetic-emotional responses we highlighted in our pilot studies (see section 4.6), which may contribute to expanding our goals. Thus the following two chapters detail two experiments conducted in contrasting setting, where we may determine the feasibility of such goals.

CHAPTER 6

Experiment 1: Naturalistic Settings

6.1 Introduction.

This first study considers the neural detection of emotion via EEG in naturalistic, real world settings. The overarching principle of this experiment is to determine the possibility of using the chosen commercial wireless EEG headset to detect and classify emotions for this purpose, whilst enlisting the AHD method. This research tests the usage of only 2 electrodes cited at position F3/F4 for Valence, and a montage of these electrodes to gauge the feasibility of detecting Arousal levels. As this is the first experiment many questions surface over the whole process, which can only be resolved through its conduction. However the key questions have been articulated and are listed as follows.

(i) **Latency**. Will a lapse in time between the viewing of a stimulus and the completion of a self -report survey, affect the participants response?
(ii) **Technology**. Is the enlisted technology appropriate, stable and robust for use in such an experiment?
(iii) **Algorithms.** Are the algorithms we have determined, for Valence and Arousal classification competent and able to return successful classification rates?
(iv) **Baseline**. Is the baseline correction method suitable for producing competent and successful classification results? Further which is the most suitable baseline correction method?
(v) **Settings.** Is the specific natural setting conducive for such an experiment, and does it allow issue free data harvesting to produce competent classification rates?
(vi) **Emotional model**. Is the dimensional model appropriate as appropriate form of classification for such an experiment?

Whilst some of the above issues may be answered through a questionnaire, other answers will arrive though processing and analysing the correlations between the self-reports and signals. Within this analysis the participants will be treated in 2 ways (i) as individuals, and secondly, (ii) as a group, by averaging their data. This has the intention of trying to determine the most successful route.

6.2. The Performance

6.2.1 The Performance: Stimulus

Figure 6.1: The publicity poster for Josephine & I [25].

As outlined earlier, a theatre performance was selected as the best potential site for conducting the first natural settings experiment (see chapter 4). It offered a real world situation, with a near perfect repeatability of a real world stimulus. Within the stimulus, we have a performer making exaggerated facial, vocal, and bodily gestures to both convey and elicit heightened emotional responses. These are enhanced by props, lighting and sound effects that operate in a precise schedule. Thus, we may consider this real world situation to be on par with a controlled laboratory type setting and stimulus.

---

[25] Retrieved from http://www.thebritishblacklist.com/wp-content/uploads/2013 /06/cush_ jumbo_josephine.jpg

Following is the related information regarding the performance.

Title : Josephine & I
written and performed by Cush Jumbo (Olivier nominated)
Directed by Phyllida Lloyd.(dir: The Iron Lady, Mamma Mia)
Bush Theatre. 7 Uxbridge Road, London. W12 8LJ (13th August- 17th August 2013)

Perfomance Synopsis :

Josephine Baker: Jazz sensation, political activist and international icon from the ragtime rhythms of St Louis and the intoxicating sounds of 1920s Paris, to present day London, Josephine and I intertwines the story of a modern day girl with that of one of the greatest, yet forgotten, stars of the 20th century.

Olivier nominated Cush Jumbo stars in her debut play, with live music and dance bringing to life the contemporary legacy of “...the most sensational woman anyone ever saw” (Hemingway). The award-winning Phyllida Lloyd directs (The Iron Lady, Mamma Mia!)

6.2.2 Performance selection:

This particular performance was selected from a choice of several potentials that were resourced via Internet searches. The selection criteria was for; a solo performance by a performer of high calibre, a reputable mid sized theatre, a continuous run of performances, and emotionally provocative content.

Being aware of Josephine Baker's story, and having read reviews of both the performance and the theatre, a dialogue ensued with the theatre, whereby permission was granted to make five recordings and one preview test recording. Further, the theatre offered assistance to ensure the experiment would run smoothly without any issues, and simultaneously not impact in anyway on either the performer or the enjoyment of a paying public audience.

6.2.3 Participant selection:

An email inviting responses to an open call for five right handed participants of both genders was sent to the residents of EECS, MAT students, and Drama researchers based at Queen Mary, University of London. In order for there to be no bias in the selection, the first five respondents were selected for the study. The study group consisted of three females and two males whose ages ranged from 25 to 34 with a mean age of 30.8 years. Four of the group comprised of PhD students based at Queen Mary, University of London, whilst the final participant was the partner of a PhD student also based at this University.

All participants had attended the theatre during the previous year. Three of the group cited they were 'occasional' visitors of Theatre performances, one cited an 'often' attendance, whilst the final participant expressed a 'rare' visitation.

6.2.4 Stimulus Clip selection:

Clip Ref	Clip Description
Clip 2	A section of dialogue, describing Josephine hanging out with her father a drummer in a bar, where she get's positive attention for her dancing.
Clip 3	An argument between her parents.
Clip 4	Her boyfriend Willy, proposes.
Clip 5	Josephine gets married, becomes pregnant, then experiences a still birth.
Clip 6	Cush talking about an audition for a show in New York.
Clip 8	Cush talks of filming in New York.
Clip 9	Josephine moves to New York.
Clip 12	Josephine leaves America and sings 'Im Sorry'.
Clip 13	Cush takes a pregnancy test, transforms into Josephine and Sings a song.

Table 6.1: A brief description of the selected clips for playback to participants in the post performance SAM test.

The researcher experienced the performance a week before the experiment. Here an audio recording and a test EEG signal were taken. The audio recording was transcribed. Careful consideration was given for the selection of 'clips' to be replayed to the participants against which they would give post-performance SAM test ratings. A range of different moods, scenarios, and actions were included in the selection, with the intention of eliciting a variant series of emotional responses.

In Josephine & I, the performer weaves through time, place, situations and characters. It was felt that coherence was key; so 'enclosed' sections of dialogue were chosen to make up the clips, which also defined their length. As it would have been subjective of the researcher to tag these clips as eliciting an expected emotion, no tags were applied and the clips were left open to interpretation by participants. Due to previous experiences of the EEG recording interface crashing during recordings in excess of 1 hour, all of the selected clips were chosen from within the first 55 minutes of the performance. Above is a brief description of the selected clips (see Table 6.1).

Whilst it was considered that the clips to be replayed to participants for the SAM test might be taken from the audio recorders they were to carry on person throughout the performance, it was felt that editing these clips 'on site' would create a large latency between the performance and the interview. Thus, the replayed clips were extracted from the initial recordings made by the researcher for all experiments. The appropriateness of this was questioned in the post-performance interview, in order to ascertain whether the recording allowed the participants to rein-visage the performance, and importantly re-experience the emotion they had experienced

6.3 Experimental Structure: Procedure

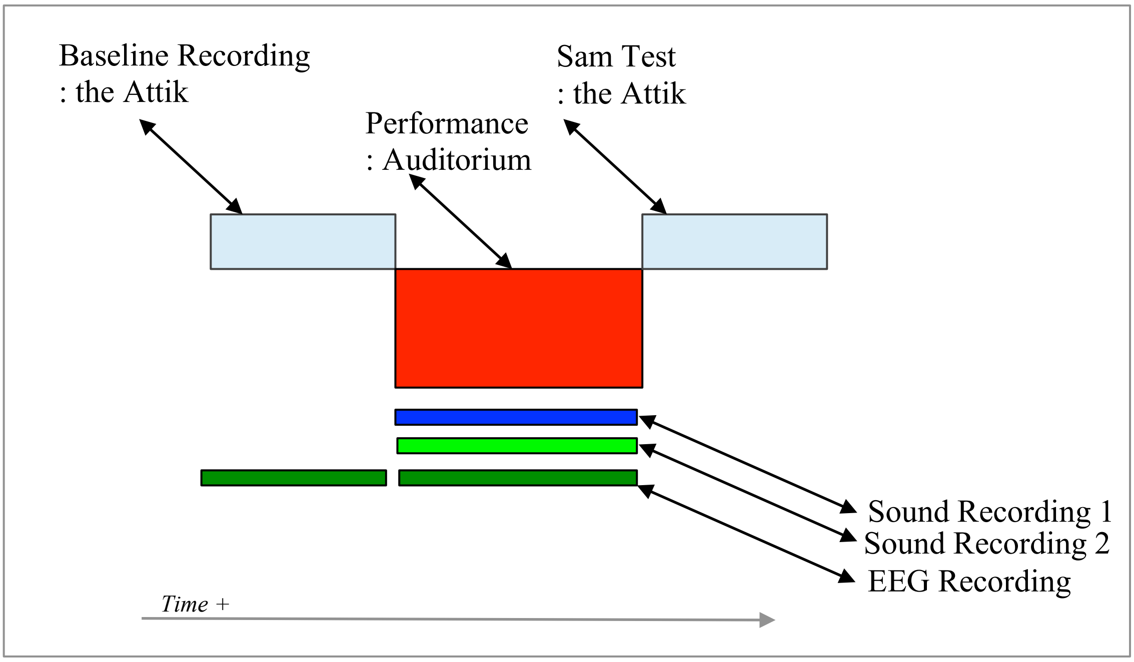


Figure 6.2 : A diagram of the experimental procedure.

The Bush Theatre provided a designated space called The Attik, as the base for the experiment. The Attik, is a large, empty, clean rehearsal space located on the top floor of the Theatre. Whilst being in direct proximity to a busy west London high street, its elevated position reduces external noise to a low amplitude therefore providing little or no distraction. Further, as the space was allocated for this projects purposes, there were no physical disturbances throughout any of the experiments. Regarding the space; the floor is of natural wood, the walls are painted white, thus there was little in the form of visual stimulation or possible distraction.

On arrival to the theatre, the participant was escorted to the Attik. Here, they were provided with the information sheet and a consent form to sign. The procedure of the experiment was explained (see Fig 6.2), and the participant was encouraged to ask questions on any factor that was not clear.

The participant was shown how to fit and remove the headset, should they wish to terminate the experiment at any time. Once they were comfortable with this process, the headset was put in place and the researcher made any required adjustments to ensure a clear signal was being received. The participant sat in a chair facing the blank wall at a distance of approx. 10ft. Having been made aware of the function of a baseline recording s/he was told to simply relax. The recording was started and the participant was left alone in the space for a maximum of 10 minutes. This process was followed for all baseline recordings.

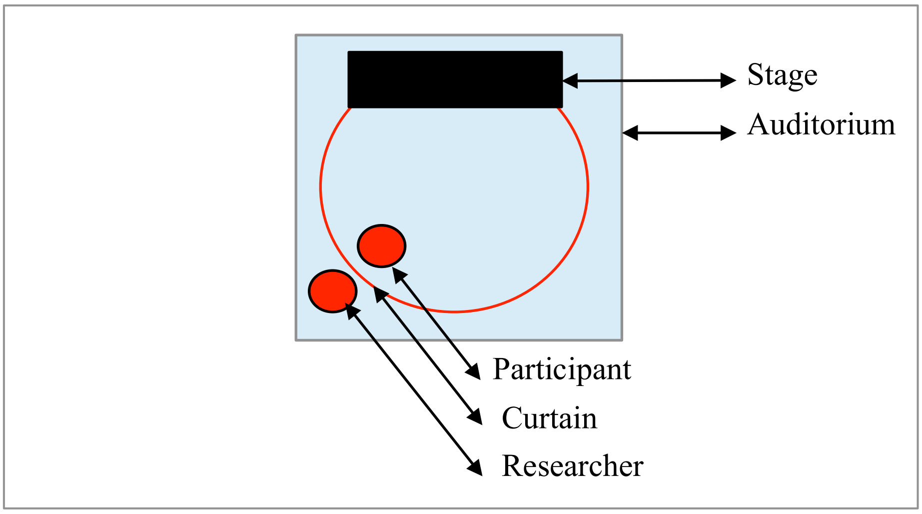


Figure 6.3: A bird’s eye view of the Theatre auditorium.

The participant was then led to the auditorium. Two audio recorders were used in this process. An apple i-pod, was the audio recorder designated for the participants sonic space. This was time-synched with the researchers apple Macbook pro that was used throughout all the experiments. The researcher used an Edirol Hdr as a second recorder for his sound space. This was mainly used to narrate the performance timings and any unexpected occurrences in the EEG signal. Once these were started in turn with audio commentary, they were left to record continuously until an arranged post-performance meet up.

As arranged by the Theatre, The Front of House manager then guided the participant to their seat; the same seat was used for all participants. From this juncture the researcher and participant would not have any interaction until after the performance. The Front of House manager then led the researcher, around the exterior of the building, through the backstage area to a provided seat and desk in the auditorium. The researcher sat directly behind the participant, separated by a curtain that enclosed the performance and audience area (see Fig 6.3). In this way, whilst the participant had been informed where the researcher was located, there was no visual, or any other form of communication throughout the performance. This served to extract the physical influence of the researcher on either the experience, or the data.

All participants wore the headset for the whole performance, which lasted for approximately 1 hour 40 min. Post performance the researcher re-established contact with the participant, In all cases the headset was removed by the participant after the performance. The participant was then escorted back to the Attik to perform the post stimulus interview and questionnaire. All participants sat at the same table in the same location for the interviews.

The nature of a SAM test and an explanation of each vector; Valence, Arousal, Dominance and their relationship to the measuring of emotion were provided. Further prompts for any required clarification were presented.

The nine selected clips, as outlined above, were placed in order on the Macbook pro screen in separate QuickTime players. Nine copies of the SAM test were provided. Firstly, a test clip was played through the headphones provided for the participants, to ensure that a comfortable sound level was achieved. Once it had been confirmed that the participants were able to cycle through the clips in the correct order, the researcher left the room for up to 20 minutes, returning only to check they were managing.

Once the participants had let it be known that this process was complete, the researcher then provided a questionnaire to be filled out (see Appendix A), and finally a receipt for their financial compensation. Each participant was given a manuscript of the performance and with this the experimental procedure was complete.

6.4 Data Processing and Feature Extraction.

Firstly for each participant, the EEG signal for each individual clip was extracted. For space saving and efficient processing, only the signals for the electrodes of interest F3/F4 were retained and exported to Matlab 2012b.

Here, the signals were passed through a bandpass filter to extract only the Alpha frequency range (8-13 Hz). This also served to eliminate/ limit the majority of potential artefacts that may affect the signals, as detailed in chapter 5 (see chapter 5). A Sliding Fast Fourier Transform (DFT) was applied at a sample rate of 1024 Hz, with a 50% overlap, giving frequency representational values of the signal every four seconds.

The Asymmetric Hemispheric Difference algorithm for Valence, and the prospective algorithm for Arousal tested throughout this research, were applied. These are as follows (see Table 6.2).

Valence algorithm	log(Alpha, Right Hemisphere) -log(Alpha Left Hemisphere)
Arousal algorithm	(Alpha Right Hemisphere + Alpha Left Hemisphere) /2

Table 6.2: The Valence and Arousal algorithms used for this experiment.

Two baseline correction values were created. A baseline (B) mean was calculated from the participant's baseline recording, and a second baseline (ACB) mean was generated through the concatenation of all a participants EEG data for the clips, and an averaged value calculated. This was conducted for each participant for both of the Valence and Arousal vectors.

For each clip and vector, the baselines (B) and (ACB) were extracted from each data point in the signal. A mean value was then calculated from the resulting series of values for each clip and vector. Finally this value was reduced to either a positive (+) or negative (-) signifier, dependent on whether the value was positive or negative. Naturally, as Alpha is seen as an inverse signal of attention, for Arousal this signifier was inverted. This process was conducted for all participants and respective signals. In turn, the survey results were also reduced to a positive (+) , Neutral (N), or negative (-), signifier dependent on their value ; >3 = (+). ==3 =(N), <3 = (-). The same procedure has been followed throughout this research (see Fig 6.4, below).

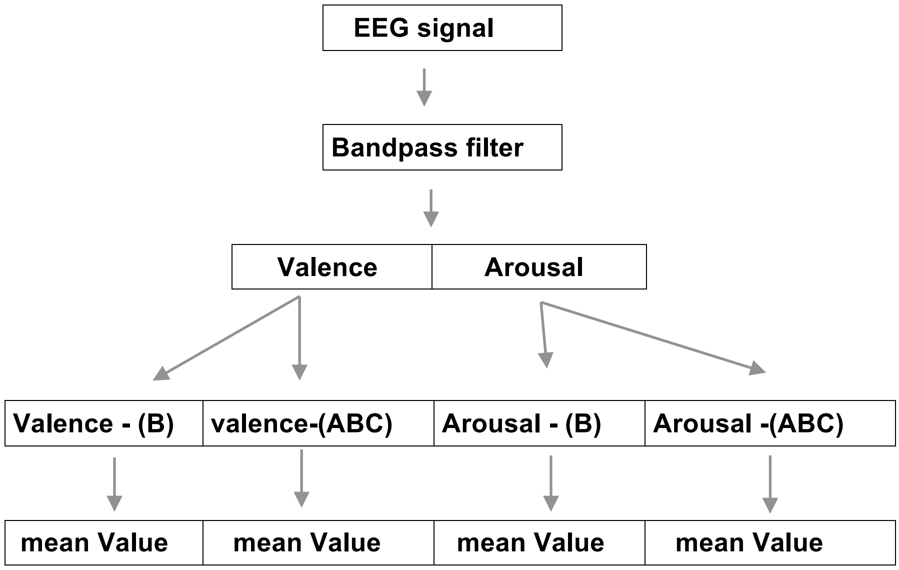


Figure 6.4: A diagram of the method from signal acquisition to output value for classification

6.5 Survey Results

6.5.1 SAM test results

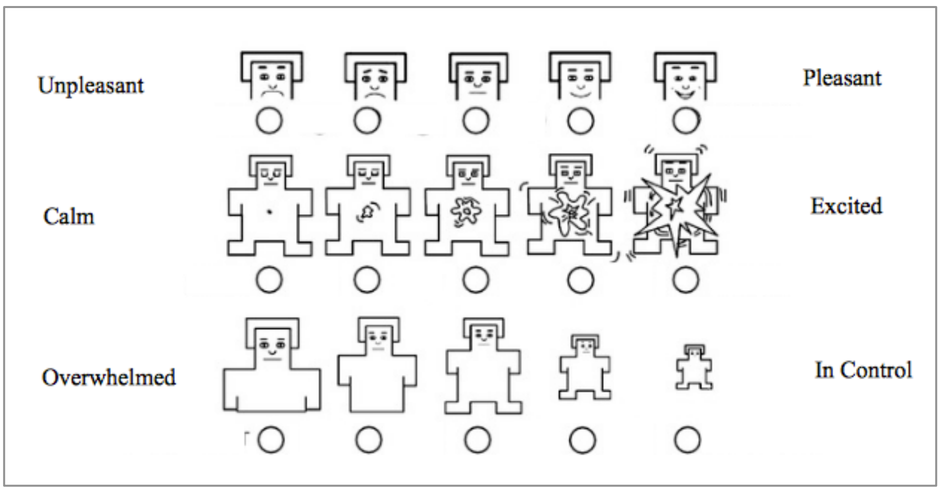


Figure 6.5: The SAM test given to all participants

Here, a summary of the SAM test responses will be presented, highlighting the most important factors. Figure 6.5, displays the SAM test used throughout this experiment. The SAM test consisted of three componential vectors which are consensually used to assess emotion dimensions of; Valence, Arousal, and Dominance. These were given on a 5-point scale of 1-5 for each. The participant was told to mark a scalar on each vector that reflected their felt emotions. Following are three tables (see tables 6.3, 6.4, 6.5) that chart the responses of all participants for each vector.

All Participant Responses for Valence							
Clip	**P1**	**P2**	**P3**	**P4**	**P5**	**Mean**	**STD**
2	2	4	3	4	5	3.6	1.14
3	1	4	2	1	3	2.2	1.3
4	2	4	3	3	4	3.2	0.84
5	1	4	2	4 >>>1	2	2.6 / 2	1.34 / 1.22
6	5	5	5	5	5	5	0
8	3	5	4	4	5	4.2	0.84
9	4	5	4	5>>>3	5	4.6 / 4.2	0.55 / 0.84
12	3	4	3	4 >>>2	4	3.6 / 3.2	0.55 / 0.84
13	2	2>>4	2	2>>>3	1	1.8 /2.4	0.45 / 1.14

Table 6.3: The participants SAM test responses for the Valence Vector. (1= unpleasant, 5 = pleasant, >> indicates a reported double value).

All Participant Responses for Arousal							
Clip	**P1**	**P2**	**P3**	**P4**	**P5**	**Mean**	**STD**
2	4	3	2	3	2	2.8	0.84
3	3	3	4	5	3	3.6	0.89
4	3	4	3	2	1	2.6	1.14
5	3	4	4	2	1	2.8	1.3
6	5	4	2	3	4	3.6	1.14
8	3	4	2	4	4	3.4	0.89
9	4	4	2	4>>>2	4	3.6 / 3.2	0.89/ 1.1
12	1	4	1	2	4	2.4	1.52
13	5	4	4	4	1	3.6	1.52

Table 6.4: The participants SAM test responses for the Arousal Vector. (1= calm, 5 = excited, >> indicates a reported double value).

All Participant Responses for Dominance							
Clip	**P1**	**P2**	**P3**	**P4**	**P5**	**Mean**	**STD**
2	2	3	4	2	4	3	1
3	2	2	3	1	3	2.2	0.84
4	4	4	5	5	5	4.6	0.55
5	5	2	3	3	4	3.4	1.14
6	5	5	5	4	3	4.4	0.89
8	3	4	5	3	4	3.8	0.84
9	1	4	4	3	4	3.2	1.3
12	5	4	4	4	4	4.2	0.45
13	3	3	1	2	4	2.6	1.14

Table 6.5: The participants SAM test response for the Dominance Vector. (1= overwhelmed, 5 = in control)

The Ranges for Each Vector Between Participants			
Clip	**Arousal range**	**Valence range**	**Dominance range**
2	*2-4*	*2-4*	*2-4*
3	*3-5*	*1-4*	*1-5*
4	*1-4*	*2-4*	*4-5*
5	*1-4*	*1-4*	*2-5*
6	*2-5*	*5*	*3-5*
8	*2-4*	*3-5*	*3-5*
9	*2-4*	*3-5*	*1-4*
12	*1-4*	*3-5*	*4-5*
13	*1-5*	*1-4*	*1-4*

Table 6.6: Circled in red are the only instances where there is a consensual agreement across the groups SAM test of whether a clip was either negative (1-2), or positive (4-5) value.

In Table 6.6, the response range across all participants for each clip and dimensional vectors are set out. It is striking to note that there is literally no overall consensus between the participant responses for all clips. Out of 27 separate fields, there are only three fields (as circled in red) of overall agreeability as to whether the vector for the clip is of either a negative (1-2), or postive value (4-5). Further, two of the three agreements are for the Dominance category, which at this stage is not a vector we are focusing on. From this we may infer that subsequent individual signal processing of the groups EEG may reflect this non-con-sensuality.

As the response range crosses the neutral point in nearly all cases it cannot be determined that the clips are neutral either. For example for clip 13 we have a Valence range of 1-4, an Arousal range of 1-5, and a Dominance range of 1-4. This almost covers the whole range of all scales.

6.5.2 Questionnaire Results.

The provided questionnaire was divided in to three subsections (see appendix A.5, for full questionnaire). The first part focused on the EEG headset technology used in the experiment and was designed to understand whether this mobile headset was appropriate for such a study.

In their responses 100% of participants agreed that the headset was both easy to fit, and comfortable to wear. Whilst all were aware of wearing the headset, none of them felt that it distracted from the performance. However all the participants agreed that towards the end of the performance it did start to become slightly uncomfortable. All the participants wore the headset for 1 hour 40 minutes. The researcher previously found that wearing the headset for more than an hour gave an uncomfortable tightness, especially in the mastoid region from the reference electrodes, and also in the temporal lobe region.

Wearing the headset in public does bring attention, and four out of five of the group were either asked questions by the public in regards to the headset, or became aware of others attention directed to them pre-performance, but not during.

The second section of the questionnaire, focused on the suitability of the sound clips used in the post-performance survey, and most importantly whether they were able to bring back the emotions experienced during the performance.

(Question) did the sound clips (a) bring back emotions you felt when you watched the performance. (b) give new emotions. (c) other

	A	B	C
PARTICIPANT 1	X		
PARTICIPANT2	X		
PARTICIPANT3	X (but not so strong 2nd time)		
PARTICIPANT4	X		
PARTICIPANT5	X	X emotions from the memory of emotions.	

Table 6.7: Participant Responses to the question of latent felt emotions

The whole group found the clips to be clear and audible, and importantly diluted any concerns of a time delay between performance and survey affecting their responses. There was a 100% consensus that upon hearing the survey sound clips they were able to both locate where in the performance this 'scene' occurred and also re-experience the same emotions they felt during the performance. One participant highlighted two answers, stating that they also gave new emotions in the form of "emotions from the memory of emotions". Their precise responses to this important question are shown in table 6.7.

(Question) did you find it easier to fill in the SAM test or to write the emotion in words?

	SAM	WRITE	COMMENTS
PARTICIPANT 1	X		
PARTICIPANT2	X		
PARTICIPANT3		X	
PARTICIPANT4	X		I couldn't think of descriptive words
PARTICIPANT5	X	X	both hard to gauge/ hard to word emotions

Table 6.8: Participant Responses to the question of the most suitable annotation method

In Table 6.8, the majority of participants demonstrate a preference for filling in variable responses in a SAM test against filling in a keyword response.

In summary of the findings from the questionnaire; whilst this was a small test group, the most important factor was that all participants felt that performing a test in such a way allowed them to re-conjure the emotions that they had felt. Whilst it has been stated (Healey, 2007) that the delay may change and alter the emotions, in this instance we found unanimous reporting that by conducting a SAM test whilst the performance is still fresh in mind and memory, we are able to extract faithful self-reports at this level of detail in relation to the stimulus, post-performance.

6.6 Results Experiment 1

6.6.1 Evaluation (i): Individual Survey to Individual Signal.

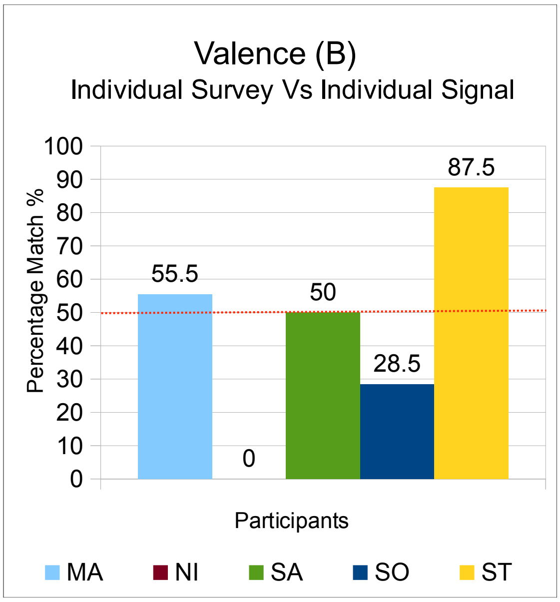


Figure 6.6: Successful classification rates Valence (B)

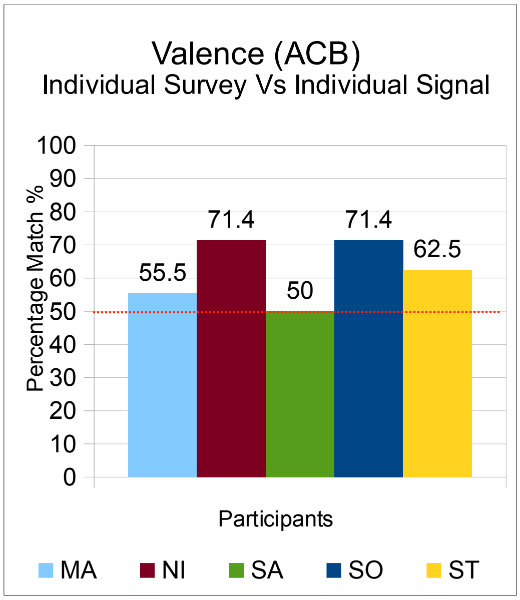


Figure 6.7: Successful classification rates Valence (ACB)

If we look at the resultant table for correct classification rates in a binary system of (+) and (-), when comparing the individual participants signal to their own individual self reports for the vectors of Valence and Arousal, with the two forms of baseline correction, (B) and (ACB), we can note: For Valence utilizing baseline (B), we find no consistent relationship between the groups individual Valence signals and self reports. Whilst one participant (ST) scored highly this is not apparent for the group where we find only 2 of the 4 participants clear the level of randomness that is 50%. A further participant is in-situ on this boundary. The highest successful classification rate is 87.5 % followed by 55.5%. There is one classification result far below random at 28.5 %. This result is for participant (SO) for whom the majority of the baseline recording was affected by unexplained interference, hence only a very small portion of this could be retained for calculating this measure. Rather than exclude this participant from the experiment, we can still use their signals for our baseline (ACB) evaluations. For participant (NI) no pre-performance baseline was taken due to time constraints so no values for baseline (B) could be incorporated. As this was the first recording the theatre wanted to ensure the participant was seated in the auditorium long before the majority of audience members arrived, and as a new set of electrodes were used the set up time took slightly longer than usual.

When we assess the classification rates for baseline (ACB) we can see that 4 of the 5 participants score higher than the random level (50%), with the remaining participant on its border, with 3 of the participants achieving above 60%. Whilst it may be tempting to state this as a significant result, this is only within population of 5 participants, and their correct classification percentage rates are not all substantially above random to provide full confidence in these results However, If a larger population was tested and achieved a consistency of above random rates, then we could be more confident.

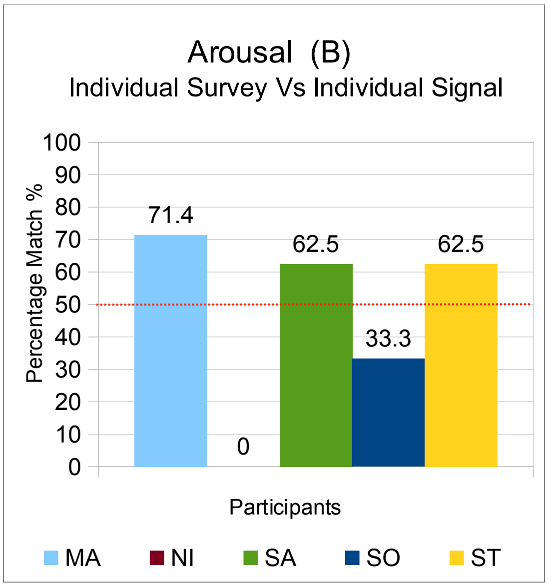


Figure 6.8 Successful classification rates: Arousal (B)

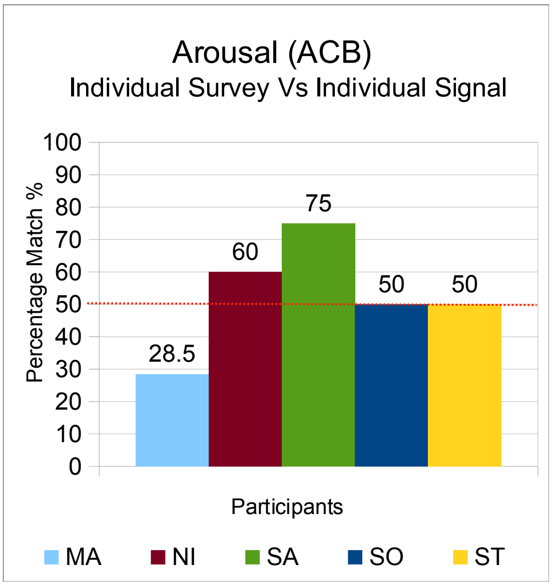


Figure 6.9 Successful classification: rates: Arousal (ACB)

In turn, for Arousal classification results, we find 3 of the 4 participants for baseline (B) register above 60%. Again it should be highlighted that participant (SO) is the participant who falls below the random level with a successful classification rate of 33.3%. These successful rates fall slightly when using the baseline method (ABC). We can view only 2 participants registering above 59.9 % successful classification, with 2 on the 50% boundary, and one below. These results display little difference from chance results, although to be sure a larger sample population would be needed. It may be of note that the highest classification rate for baseline (B) becomes the lowest for baseline (ACB). The classification rates for both Valence and Arousal are gathered in table 6.9

Classification % : Individual Signal versus Individual Survey				
Participant	**Valence (B)**	**Valence (ACB)**	**Arousal (B)**	**Arousal (ACB)**
SA	50%	50%	62.5%	75%
ST	87.5%	62.5%	62.5%	50%
MA	55.5%	55.6%	71.4%	28.5%
NI	-	62.5%	-	60%
SO	28.5%	71.4%	33.3%	50 %

Table 6.9: Successful Valence and Arousal classification % rates for all participants

r-Correlation: Individual Signal means and Individual Surveys								
	Valence (B)		**Valence (ACB)**		**Arousal (B)**		**Arousal (ACB)**	
Participant	*R*	*P-Value*	*R*	*P-Value*	*R*	*P-Value*	*R*	*P-Value*
SA	-0.09	*0.81*	*-0.09*	*0.81*	-0.67	*0.05*	-0.67	*0.05*
ST	0.41	*0.28*	*0.41*	*0.28*	-0.15	*0.69*	-0.15	*0.69*
SO	0.37	*0.36*	*0.37*	*0.36*	0.48	*0.23*	0.48	*0.23*
MA	0.05	*0.89*	*0.05*	*0.89*	0.1	*0.81*	0.1	*0.81*
NI			*-0.01*	*0.97*			0.17	*0.67*

Table 6.10: Individual participants Correlation Coefficient r, was calculated between their Individual Signal (mean) and Individual Surveys for all clips. Each analysis used 9 datapoints.

A consideration was made as to whether there may be some form of direct linear relationship between participants SAM test ratings and their averaged signal values. We evaluated the means of the data signals, against the SAM ratings, which gave 9 data points for each participant for the 9 clips. In table 6.10 we have calculated the Correlation Coefficient *r,* on case-by-case basis, from which we can clearly view that there is a very weak linear relationship between them. The only marginal relationship can be seen for Participant SA, Arousal (B) & (ACB), ($r = -67$, $n = 9$, $p = 0.05$).

6.6.2 Evaluation (ii): Group Survey to Individual Signal

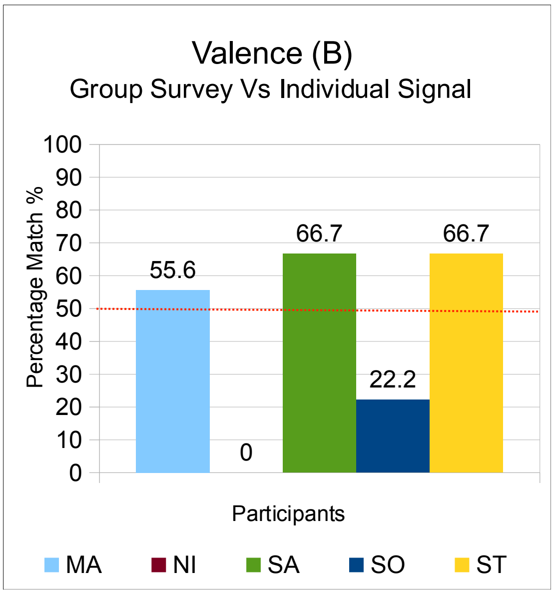


Figure 6.10 Successful classification rates: Valence (B)

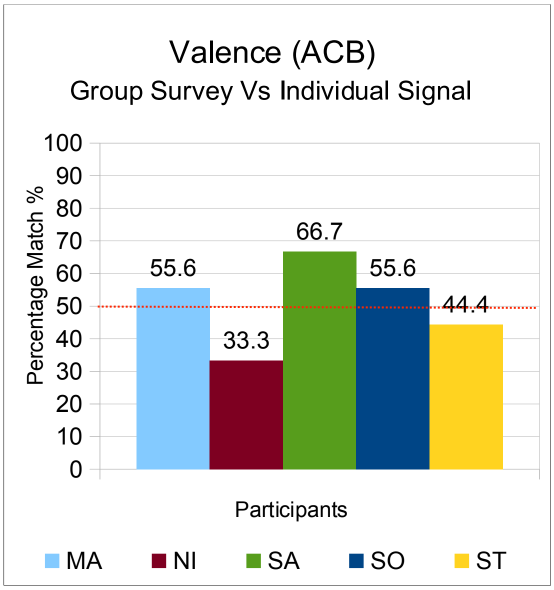


Figure 6.11 Successful classification rates: Valence (ACB)

For the second in the series of evaluating correspondences between Survey and Signal, we have taken a mean value of the groups SAM responses for each clip and compared this against the signals of the individual participants. In this way we can assess whether by further averaging their results we may be able to increase successful classification to a given stimulus, and whether this is a more appropriate method.

As revealed in figure 6.10, for baseline (B) we have the same pattern of the 3 same participants achieving above the random level. These results are a slight improvement on the previous results (individual survey to individual signal baseline (B), see Fig 6.6) with respective results of; 55%, 66.7% and 66.7 %. We also again find that participant (SO) scores well below the random indication level, and the rest of the group with a classification rate of just 22.2%.

In figure 6.11, we can see the same comparison results for baseline (ACB). Here we can view a reduction in successful detection rates, compared to our previous (ACB) comparison (see Fig 6.7). Here only 3 of 5 participants score above 55% with the highest individual participant correct classification percentage of 66.7%.

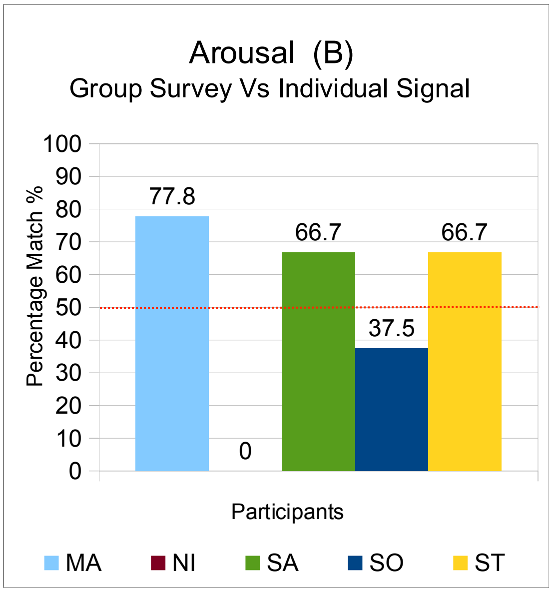


Figure 6.12 Successful classification rates: Arousal (B)

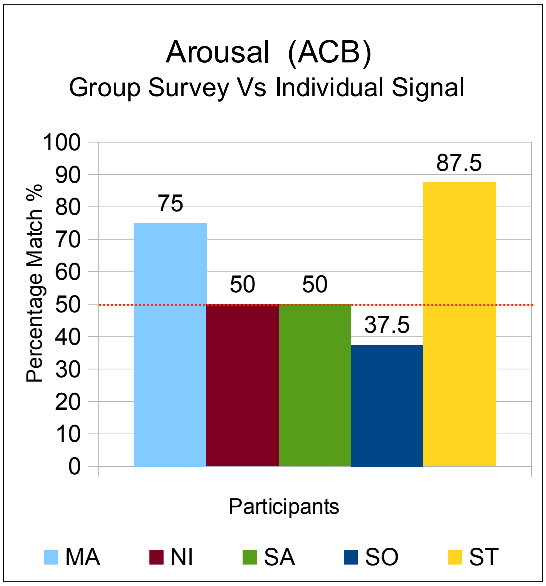


Figure 6.13 Successful classification rates: Arousal (ACB)

For Arousal we can note a greater successful classification rate using this analysis combination. 3 of the group of 4, score more than 65% for baseline (B), again participant (SO) scores the lowest. For baseline (ACB) we can see the highest single correlation for successful Arousal correlation at 87.5%, with another participant at 75%, and 2 further participants on the random boundary with 50%, and 1 below (SO). Whilst the results have improved for Arousal detection, the inverse is true for Valence. Table 6.11 presents the classification results numerically.

Classification % : Individual Signal versus Group Survey (mean)

Participant	Valence (B)	Valence (ACB)	Arousal (B)	Arousal (ACB)
SA	66.7%	66.7%	66.7%	50%
ST	66.7%	44.4%	66.7%	87.5%
MA	55.65	55.6%	77.8%	75%
NI		33.3%		50%
SO	22.2%	55.6%	37.5%	37.5 %

Table 6.11 Successful Valence and Arousal classification % rates for all participants

6.6.3 Evaluation (iii): Group Survey to Group Signal.

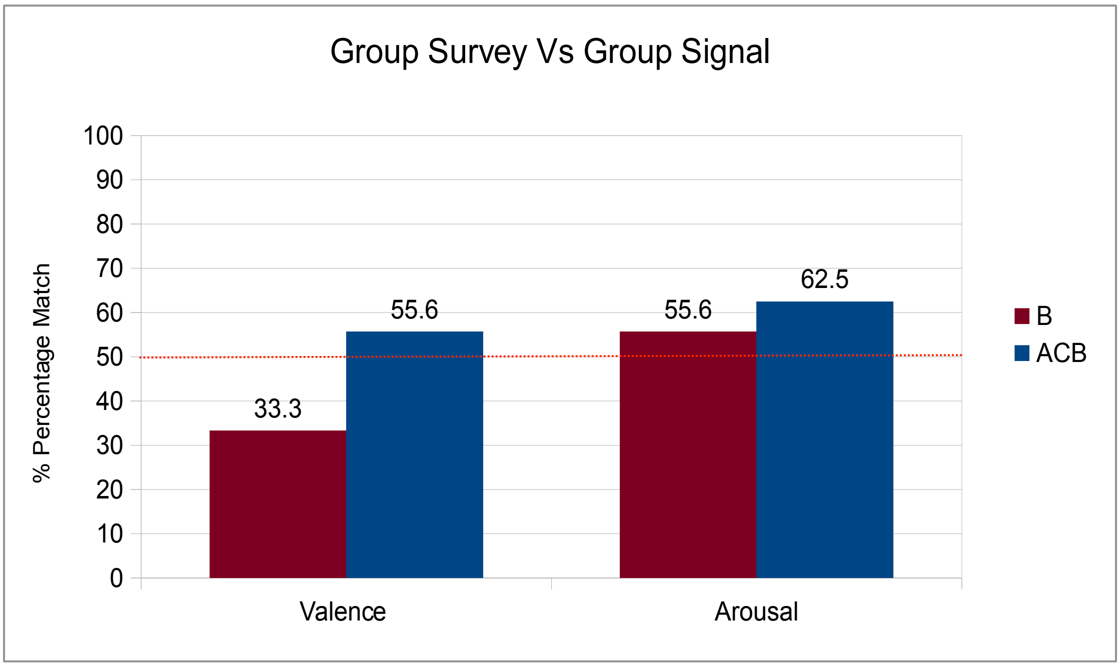


Figure 6.14 Successful group classification rate: group survey versus group signal

In a final consideration, both the participants EEG signals and surveys, were reduced to a single entity by taking a mean representational value for both. In terms of Valence, for baseline (B), we have a result that falls below the random level of 50%, with a successful classification rate of just 33.3 %. This rises slightly above random when using baseline (ACB) to 55.6%.

For Arousal we do find more successful results. Baseline (B) produces a success rate above 55%, whilst with baseline (ACB) we arrive at the most successful measure of this comparison combination of 62.5%.

r-Correlation : Group Signal (abs mean) and Group Survey

	Valence (B)	**Valence (ACB)**	**Arousal (B)**	**Arousal (ACB)**
r Value	0.08	0.25	-0.32	-0.35
p-Value	*0.83*	*0.51*	*0.39*	*-0.35*

Table 6.12: Group Correlation Coefficient r, between the group signal (mean) and survey (mean) for each clip, the 9 clips provided 9 data-points in each evaluation.

As with previous considerations, in this form evaluation no evidence of a liner relationship between the signals and survey result could be determined (see table 6.12). Here the nine mean values of the participant's signal value for each clip were considered alongside the groups nine mean survey values for the corresponding clip.

6.6.4 Results Summary

Thus, when considering the 3 types of evaluation combinations we have performed, we may state that the best results when considering the groups EEG signals individually for Valence were; using baseline (ACB), evaluation (i) which contrasted Individual signals to their Individual surveys. Here, 4 of the 5 participants scored above 55%, with the final participant on the random level indicator of 50 %. For baseline (B) the best Valence results were achieved for evaluation (ii), when comparing the individual signals to the averaged groups surveys. Here, if we were to discard participant (SO) due to the issues of a problematic baseline, we would find a 100% rate of individuals having classification results higher the 55%. Yet with a population of only 3 participants, this may not be deemed significant, as such a statement would require a larger participant group. In both of these best results, had all the percentages resulted as significantly high, for example close to 75% then we may have reported with more confidence.

When both Valence EEG signals and SAM test survey were averaged to function as group reading in evaluation (iii), of the two baselines, baseline (ACB) was the only one to register above random levels at 55.6%.

For Arousal the best classification results were achieved for baseline (B) in both evaluations combinations between Individual signal to Individual survey, and Individual signal to Group survey. If we discard participant (SO) from our results, we find the whole population score above 62% successful classification rates in both instances. This outperforms Arousal results for baseline (ACB) where we find only 2 participants scoring above 59.9%, and 2 further participants on the random level of 50% for both evaluations. Of the two instances, the second evaluation (ii) combination was best.

For the averaged Arousal signals in our third evaluation we achieved success rates above 55.6% in both instances with baseline (ACB) the best performer at 62.5%.

From all our assessments of a linear relationship between the signal and survey values we were only able to determine a very weak correlation.

6.7 Conclusions and Discussion.

In this primary study we have conducted an experiment into whether it is possible to determine emotion experiences via EEG in naturalistic type settings. Naturally with a first study, the whole process and its methods are under review. However a few prominent questions were highlighted at the outset of the experiment. These are again listed below:

(i) **Latency,** Will a lapse in time between the viewing of a stimulus and the completion of a self -report survey, affect the participants response?
(ii) **Technology,** is the enlisted technology appropriate, stable and robust for use in such an experiment?
(iii) **Algorithms.** Are the algorithms we have determined, for Valence and Arousal classification competent and able to return successful classification rates?
(iv) **Baseline**. Is the baseline correction method suitable for producing competent and successful classification results? Further which is the most suitable correction method (B) or (ACB)?
(v) **Settings.** Is the natural setting conducive for such a experiment, and does it allow issue free data harvesting to produce competent classification rates?
(vi) **Emotional mode**l. Is the dimensional model appropriate as appropriate form of classification for such an experiment?

In regards to the question of latency between experience and report, the participants experienced a continuous performance lasting approximately 1 hour 40 minutes. Following this the participants were led to a designated space where they listened to nine short audio clip extracts of the performance, and completed a SAM test survey for each. In the questionnaire they were asked whether they could remember the sequential occurrence of the clip in the performance, and most importantly whether this brought back the emotions they had experienced. There was a unanimous agreement that they were able to locate the clips juncture in the performance, and significantly that they were able to re-experience the same emotions they felt during the performance.

This suggests the suitability of this framework, and opens up the potential for such studies to be conducted 'in the wild'. This is without the need for an interruption of stimulus to fill in forms of self-report. Further this suggests that such a stimulus need not be limited in time and scale, and can be part of a larger context. This also directs that it may be possible to cautiously move beyond a theatre setting into further diverse settings. Naturally for group studies, there is the formal question of repeatability, but as shown above with careful consideration this may be achieved to a good degree. This finding also presents the grounding for longitudinal studies (without repeatability) for a single participant over many carefully considered diverse stimulus settings.

This study also asked, whether the enlisted portable technology was robust and appropriate for such an experiment. All participant reports objectively confirmed that the EEG headset was suitable. It was comfortable to wear, unobtrusive, with a simple and minimal set up time. The technology was also deemed robust by the experimenter, the EEG signal was maintained throughout the recordings, and the specialized recording interface Testbench was stable throughout.

In this experiment the two emotional dimensions of Valence and Arousal, were explored. For Valence the repeated peer reviewed method, of Asymmetric Hemispheric Difference was tested. This incorporated using a minimal portable set up, of only 2 electrodes F3/F4 for Valence detection. Their montage into a single electrode was used to measure Arousal in a speculative approach, to consider whether Alpha Spectral Power may be a measure of this.

The classification results for Valence when comparing an Individuals EEG signal to their self-report for this measure are indicative of a marginal relationship. Utilizing the baseline correction method (ACB) (evaluation (i)), 4 of 5 participants had classification percentages of 55%, which are above the level of randomness of 50%. The groups respective scores of; 50%, 55.6%, 62.5, 62.5%, 71.4%, which produce an average of 60.4% are reflective of this marginal relationship. This was the most successful method tested. When we treated both signal and self-report value as single groups, by averaging both to respective single variables we were only able to achieve a classification rate of 55.6%. In future studies it would be beneficial to continue to consider the population in both ways.

In terms of Arousal, as mentioned prior there is no successful peer reviewed method. Hans Berger the inventor of EEG demonstrated through rigorous experiment that the Alpha frequency (8-13 Hz) is an inverse signature of attention. In this study we attempted to decipher whether there may be any correlation of neural activity levels measured by Alpha (8-13 Hz) in the frontal cortex region and self-reports of Arousal in a SAM test. Arousal was calculated as the averaged sum of Alpha derived from electrodes F3/F4. This was to gauge whether we may in some way be able to map this EEG measure to Arousal in a dimensional model of emotion.

For Arousal the best classification results were achieved for baseline (B) in instances of both individual signal to individual survey, and individual signal to group survey (mean). This is the case when we discard participant (SO) from the study due to the issues mentioned with baseline (B). In evaluation (ii) All 3 of the 3 participants scored in excess of 65%; with 66.7%, 66.7%, and 77.8% which produces an average score of 70.4%. For evaluation (i) success rates were 62.5%, 62.5%, 71.4%. These rates may be considered as having some marginal significance, but would require a larger participant test group, with consistent further higher successful rates to provide the necessary confidence in this measure.

The best Arousal results achieved for all 5 participants were in evaluation (ii) when comparing the Individual signals against the averaging of the groups self-report using baseline (ACB). The results returned classification rates of; 37.5%, 50%, 50%, 75%, and 87.5%. Arousal outperformed Valence classification when both signal and survey were treated as a singular entity, with 62.5% against 55.6%.

Whilst it may be tempting to use these statistics (baseline B) to present an argument for the Arousal measure, the small population incorporated in this study leave the sense that this relationship would require further investigation, and thus should be regarded as inconclusive.

It may be stated that the level of Arousal one may sense is different and notoriously more difficult to gauge than Valence. Whilst Valence may easily be deciphered in a binary consideration, Arousal may be more individualised and contextualised to immediately prior sensations, and further to a more personalised sense of history, than Valence. This was reflected in the surveys.

Two baseline correction methods were tested where firstly (B); a mean correction value calculated from the participant's baseline recording, and secondly (ACB); a mean correction value calculated from all the data for the relevant clips. Whilst baseline (ACB) outperformed baseline (B) for Valence, we found the inverse for

Arousal. Thus it is felt important to continue to include both baseline methods in any subsequent studies. As we may have noted with participants (SO) and (NI) where problems arose with their baseline recordings, we were still able to include them in this study through the second baseline correction method. It may be that for particular experimental contexts, one of the different correction methods may be more suitable than the other.

Regarding the important aspect of settings. We can confirm that there were no issues in data harvesting, and further that such a setting is very conducive for our intentions. The participants were highly engaged with the stimulus, and all reported back how involved they felt with the stimulus, which is suggestive of high levels of emotional investment and elicitation. Yet, it remains unclear whether the setting has impacted on our signals, and whether the freedom of natural movement in conjunction with a complex stimulus may have affected our classification rates.

The final pre-experiment outlined question centred on the appropriateness of the dimension model as a basis for classification. It was confirmed through the questionnaire that the participants found it easier to give dimensional values than single keywords to articulate their felt experiences.

Questions that emerged through the processing of data, considered how best to treat the participant population; either as a group or as individuals. Through the analysis we arrived at variant results. In the case of Valence the best results pointed towards the individual, whilst for Arousal these were skewed towards a group treatment, thus it may be profitable to consider both configurations in any subsequent studies. Due to economical and logistical constraints only a small population was used in this study, and this may play also play a factor in the indeterminacy of the results. As can be seen with participant (SO), in small populations, a single participant has the potential to considerably affect results. Whilst above this was for a insufficient baseline, this can also be true for other factors such as the unawareness or difficultly of articulation of labelling emotion, or simple non-engagement. All of these may affect the filling in of surveys. Thus the necessity for larger groups in future studies was noted. In the instance of longitudinal studies, a larger number of studies may serve as a resolution.

Regarding the SAM test surveys, for this study a 5-point scalar was used for each dimension. It was felt that a neutral value occurred more often than anticipated in the responses. To negate this, a larger 9-point scalar could be used to present a greater freedom of response for the participants around the neutral registration.

Whilst it is important to acknowledge some marginal degree of success with our detection results a secondary experiment needs be conducted in laboratory conditions with a larger population. Here, any questions of ambiguity towards the results may be addressed. By reducing the complexity of stimulus to controlled settings, we may be able to further understand the results of this study, and whether measuring EEG in natural setting becomes more problematic due to the possible introduction of further artefacts which may disguise the signal. This is especially pertinent due to our set up which only uses two electrodes to gather data, and further the use of a low-cost mobile commercial headset, whereas traditional EEG set-ups use clinical technology with larger electrode configurations.

CHAPTER 7

Experiment 2: Laboratory Settings

7.1 Introduction.

Following Experiment 1, which was conducted in natural settings, it became important to conduct a second contrasting experiment under controlled laboratory conditions. This would allow a better understanding of the limitations of the 'in the wild study', and whether it is because of difficult recording conditions or because of more fundamental difficulties in linking EEG to fine-grained emotional responses to complex time-based artistic materials. Here any unknowns which may arise in a real world setting and have some influence on the EEG signal and the participants responses can be eliminated or minimised, and factors thought to be important to the outcome can be isolated. Therefore this management of conditions may present an enhanced confidence in the reliability of the received data and results. Thus for this experiment all variables except the presentation order of films clips were as precise as possible for all participants. Finally the tight structure of the experiment also aids its potential duplication and repeatability. The key questions asked in this for experiment are as follows

(i) Valence, is the AHD method for detection reliable?
(ii) Arousal is the ASP method for detection reliable? Is there any basis for power levels in the alpha region being indicative of Arousal?
(iii) Will having a larger experimental **population** allow successful classification rates for both individuals and as a group?
(iv) Is the **technology** robust and reliable for this form of experiment?

It should also be highlighted at the outset of this experiment that it is questioned whether such a setting may have an impact on the Arousal classification results. As detailed in the literature review, experiments conducted in laboratory settings may decrease participant's levels of emotional investments and engagement. This is particularly relevant when using film clips. Film has a natural structure for increasing emotional investment over time, as relationships are built to the protagonists. When the clips are removed from this natural structure we may lose the power of this device. This may also be amplified with the start-stop nature of the experiment, where the participant is reminded of being in experimental conditions.

This experiment was conducted in 2 stages. Stage 1 comprised of a group of 10 participants, who would define and tag a set of 16 film clips with both an emotional keyword, and a numerical value for both Valence and Arousal vectors via a SAM test. Stage 2 comprised of a second group who under laboratory conditions would view the objectively selected clips whilst wearing a commercial EEG headset. They would also complete the same SAM test.

It may be possible that the result from both experiments can be used to support or destabilise the findings of the others validity.

7.2 Participant Selection.

An open email invitation was sent to all EECS & MAT residents, based at Queen Mary University of London. This open call requested for 10 participants (right handed) to take part in study which gauged emotional responses to film clips. The call made no exclusive distinctions between race, gender, age, position, or nationality for inclusion. To eliminate any form of bias the first 10 responses were selected as the experiments participant Group 1 (G1). Due to the volume of responses the next 10 responses were selected to take part in the second stage of the experiment to comprise Group 2 (G2).

G1's demographics were as follows; 4 Males, 6 Female, Age ranges were 24-48, with a mean age 32.4 (*std* = 7.8). Participating Nationalities were: 1 Irish, 3 British, 2 Central European, 3 Southeast Asia, and 1 America. 9 of the group were QMUL students, and 1 a QMUL staff member.

G2 comprised of 4 males, and 5 females, with an age range between 22 & 28, with a mean age of 25.4 years. Group 2's participating nationalities consisted of 1 British, 2 Indian,1 Irish, 2 Eastern European, 2 Chinese, and 1 South American . The group comprised solely of PhD candidates based at QMUL. Two original participants were excluded from the study due to the volume and quality of their hair, which prevented electrode contact with the scalp and the detection of the EEG signal. These two participants were replaced with one other to create a group of 9 participants instead of the intended 10.

In the above demographics, whilst we are selecting within the narrow population of students from a single university, we can note that we have a great diversity in terms of participating nationalities.

7.3 Experiment 2 : Stage 1

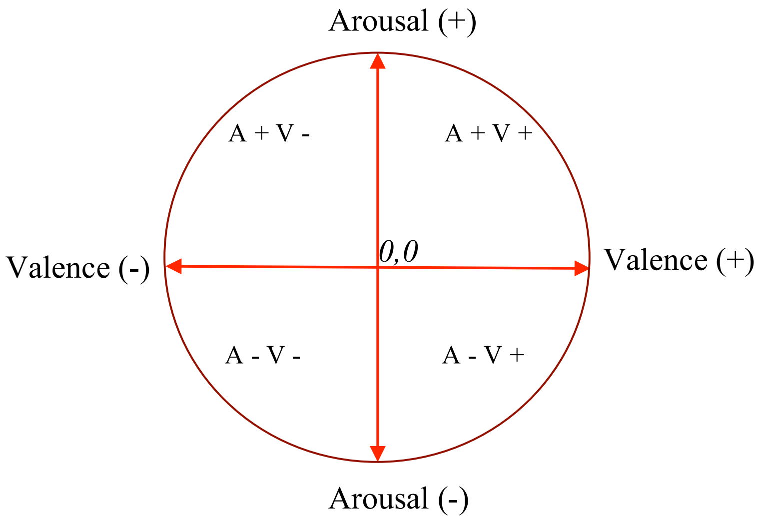


Figure 7.1, The Valence and Arousal dimensional vectors as binary conditions (positive and negative) in a circumplex space.

The intention of Experiment 2: Stage 1, was to create an objective set of variant film clips tagged with a dominant emotional keyword, and Valence and Arousal values through the means of a survey, by Group 1 (G1). Once obtained, the objectively Tagged videos would then become a stimulus set for Group 2 (G2), who would observe the chosen video clips whilst wearing EEG headsets, and also produce a second set of survey data. The analysis of the results would compare the EEG signals against both sets of surveys.

In the design of this experiment, a method published by Soleymani, Pantic, & Pun (2012) was used as a reference for constructing such a set of stimulus videos. In their experiment they used EEG detection as part of a multi-modal approach of assessing emotional responses to film clips, and for their purposes segmented the vectors of Valence and Arousal into 3 discrete sections (calm, neutral, positive). In this experiment where we are using only the single modality of EEG, both vectors; Valence and Arousal, are being divided in to 2 discrete categories Positive (+) and Negative (+) as is shown in figure (7.1).

7.3.1 Preparatory Film Clip selection

More than 100 short film clips were extracted from 60 commercial films available in the QMUL library catalogue. This assortment was edited down to a selection of 38 clips that ranged in length between 1 and 5 minutes. These clips covered a range of moods, tones, imagery, situations and actions. It was hoped that between them they would be able to elicit a variant range of emotions, from which a smaller group of objectively tagged film clips could be obtained.

The criterion for each clip was for a 'closed' segment or scene that would present a clear and cohesive situation. Adobe Final cut pro was used to edit and prepare these. The 38 film Clips were then randomised and placed on 10 DVDs.

7.3.2 Procedure.

Each Participant of G1 was offered the option of either conducting the experiment within a designated space (MAT Computer Suite, QMUL) or alternatively, to 'take-away' the experiment to be completed at their own convenience.

The 'take-away' stipulations were; that they were to watch the clips alone in a distraction free space, to wear headphones if applicable, and to watch the films in any order in time chunks of no less than 15 minutes. This is reflective of Soleymani et. al (2012) methods whereby film clips were presented through an online portal, with no designated singular point of access.

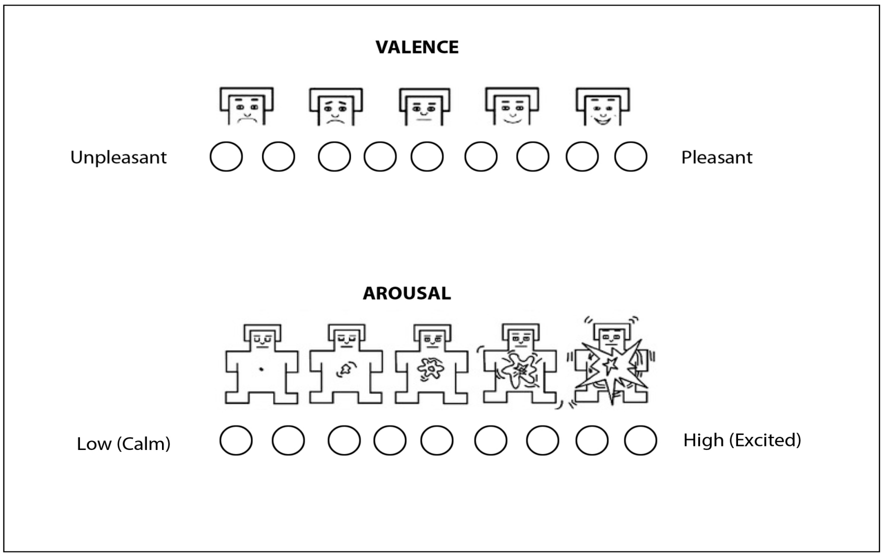


Figure 7.2: The SAM test for experiment 2 comprised of a 9-point scale on each dimension of Valence and Arousal.

Upon arrival, to the laboratory space, each participant (G1) was provided with an information sheet and verbally informed of the nature and conduct of the experiment (see appendix B; B1-B7, for all information sheets, questionnaires, and SAM tests relating to Experiment 2). An explanation of the SAM test and its variables of Valence and Arousal were provided (see Fig 7.2). The function of the keyword responses to be reported as felt emotions, and considered emotions were also relayed. This segmentation between felt and considered emotions was provided to minimise any confusion of response, and to ensure that a felt emotion was reported in each instance. Once the participants verified they understood all aspects, they signed the consent form and took the experiment package away. This consisted of a DVD containing 38 QuickTime movies, 38 SAM tests, and a written account of the verbal instructions.

From group 1 (G1), only 1 participant chose to conduct the experiment 'Live'. This experiment took place in a continuous 2-hour time block inclusive of breaks, at the MAT computer suite. 1 participant’s DVD failed to operate, thus the clips were then made available for download online via Dropbox.

Upon the return of the completed experiment participants were given a compensation payment for their time.

7.3.3 Results

Clip	Emotional Tags	Valence (Mean/STD)	Arousal (Mean/STD)
001	CON 4, JOY 3	7.2 / 1.50	2.7 / 1.68
002	JOY 4, CON 3	6.9 / 1.29	4.5/ 1.65
003	JOY 6, AMU 3	7.8 / 1.5	5.9 / 2.8
004	ANX 5, DIS 3	1.8 / 0.6	6.4 / 1.8
005	HAPPY 7, JOY 3	7.6 / 1.2	6.4 / 1.0
007	AMU 8, NEU 2	5.8 / 1.8	4.4 / 1.6
008	ANX 5, AFR 2	3.9 / 2.0	6.4 / 1.5
009	ANX 5, AFR 2	2.9 / 1.8	8.0/ 0.9
011	DIS 6, AFR 1	1.4 / 0.7	6.5 / 1.5
012	SAD 5, NEU 3	5.0 / 1.5	3.5 / 1.2
013	AMU 4, NEU 3	6.8 / 1.2	2.9 / 2.0
016	HAP 5, NEU 3	6.8 / 1.5	3.1 / 1.7
017	ANX 6, NEU 3	3.6 / 1.3	6.9 / 2.7
018	ANX 6, AFR 1	4.2 / 1.5	6.9 / 2.2
019	ANX 5, AFR 2	3.9 / 1.4	6.6/ 2.5
020	ANX 5, HAP 1	4.6 / 2.5	6.9 / 1.5
021	SAD 5, ANX 3	2.6 / 0.7	5.8 / 1.3
023	ANX 3, HAP/SAD/AFR 2	3.9 / 1.9	6.1 / 1.6
024	ANX 4, AFR 3	3.4/ 1.1	6.4 / 1.2
025	NEU 6, AMU 3	6.0/ 1.3	3.5/ 1.2
026	CON 5, HAP 3	7.0 / 1.3	2.1 / 1.0
027	HAP 3, JOY 2, NEU 2	7.6 / 1.5	3.4 / 2.5
029	AMU 8, HAP/NUE 1	7.7/ 1.3	4.6 / 2.0
031	HAP 4, CON/ NEU 2	6.7 / 0.9	3.4 / 1.8
034	JOY 5, NEU 3	6.4 / 1.2	4.5 / 2.1
036	SAD 4 , NEU 3	3.6 / 1.1	3.4 / 1.7
038	SAD 4, AMU 3	5.3 / 1.4	3.3 / 2.0
042	AMU 5, HAP 2	8.0 / 1.2	6.3/ 1.4
043	SAD 3, AFR 3	1.8 / 1.0	6.7 / 1.4
050	ANX 4, AMU 2	3.8 / 1.0	6 / 1
055	ANX 4 / NEU 4	4.5 / 4.4	4.4 / 2.2
056	AMU 4, JOY 3	7.3 / 1.2	6.6 / 1.3
057	AMU 7, CON 1	7/ 1.3	4.7 / 2.3
058	NEUT 4, HAP 2	5.7 / 0.9	3 / 1.1
059	AMU 9, JOY 1	7.3 / 1.3	4.5 / 2.5
060	AFR 4, ANX 3, SAD 3	2.3/ 0.9	6.6/ 1.6
061	AMU 5	6.0/ 2.2	3.5 / 1.7

Table 7.1: Group 1's survey responses: Column 1 is the clip reference. Column 2 shows the collated keyword responses, columns 3 & 4; Valence and Arousal mean/std values for the group. The abbreviations are afraid (AFR), amused (AMU), anxious (ANX), content (CON), disgust (DIS), happy (HAP), joy (JOY), neutral (NEU), sad (SAD).

The returned surveys were examined in a number of ways. Firstly the data was collated into a single table for the group against each film clip (see table 7.1). For each clip the keywords for felt emotions were counted, and the mean/standard deviation (std) values for Valence and Arousal across participants calculated. It is notable that there is no occurrence of a single reported felt emotion keyword for any clip across the group. Following Soleymani et. al (2012) the keyword with the highest number of registrations was attached to the film clip.

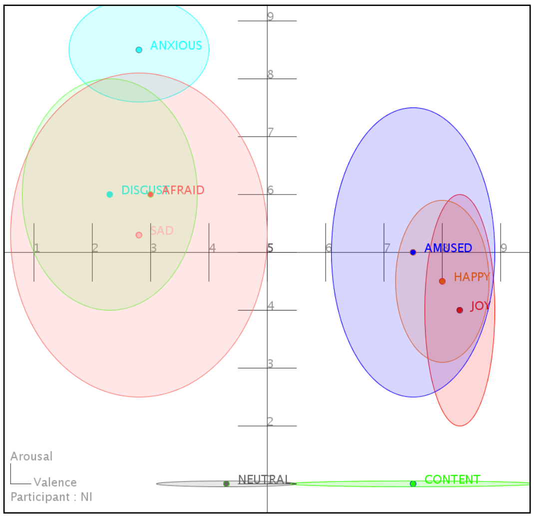


Figure 7.3: The range of responses each emotion for Participant NI

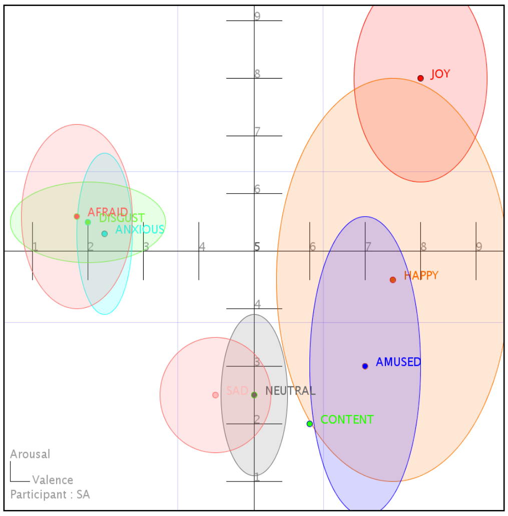


Figure 7.4: The range of responses for for each emotion for Participant SA

Secondly the individual responses of the participants were considered. For each participant the mean and standard deviation values were calculated for each reported felt emotion across all 38 clips. These were visually plotted for each individual. Figures 7.3 and 7.4, provide examples from participants (NI) and (SA). Here, each felt emotional keyword is presented next to its mean value, which is represented by a dot of the same colour. The coloured circle which extends from this represents the extent of the standard deviation, the horizontal stretch represents Valence, whilst the vertical represents Arousal. From these 2 examples, we can view the pattern of difference that was consistent throughout the group; there was not a single static value on either axis associated with any form of felt emotion, rather these are seemingly dynamic variables.

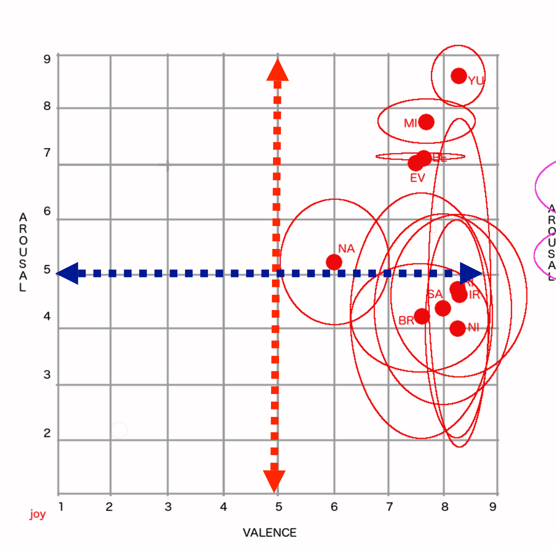


Figure 7.5: All participant responses range for Joy

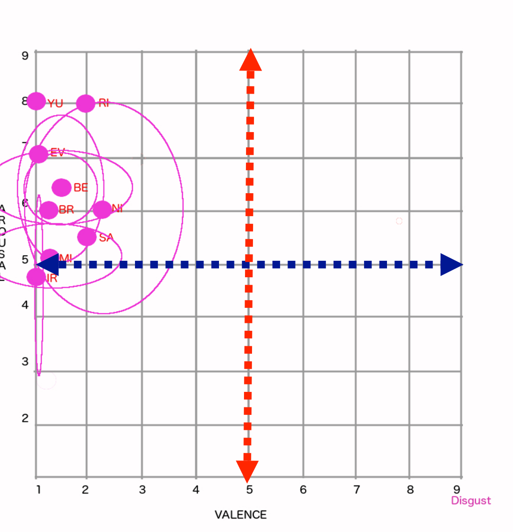


Figure 7.6: All Participant response ranges for Disgust.

N.B In both figures, on each vector >5 = (+), == 5 =(N), <5 = (-)

In a secondary consideration, the individual participant responses for each felt emotion were visually composited together. In this way it becomes possible to clearly see the inter-participant correlations for each felt emotion. Figures 7.5, and 7.6, present examples for Joy and Disgust. Whilst there was a consistent strong correlation for Valence responses, which resided within either a positive or negative space for all emotions, for Arousal there was a wider dispersion across the division between positive and negative registration. However, this was not the case for the felt emotions of Disgust, Anxious, and Afraid that do share the same Arousal space. Figure 7.5 shows the example for Joy where we have a correspondence of (+) Valence, whilst a dispersal across both (+) and (-) sections for arousal. Figure 7.6, presents the example of disgust, which again has a correspondence for valence, but also this time for Arousal. This leads to an expectation that through this study there maybe a likelihood of greater success rate in classifying Valence than Arousal.

7.3.4 Final Selection of Film Clips.

As outlined above, the prominent keyword for each clip was attached to it. Thus a selection of 16 film clips was made based on this objective data. 2 clips were chosen for each of the following emotions: Afraid, Amused, Anxious, Content, Disgust, Happy, Joy, Sad. The final list of selected films is as follows. (see Table 7.2, below).

Clip Code	**Emotional Labels** Valence	Arousal	**Dominant Emotion Tag**	**Video Clips Sources**
001	+	-	Content	The Tree of Life (Start)
003	+	+	Joy	The Holiday (House)
004	-	-	Disgust	Zero Dark Thirty (Water Torture)
005	+	-	Happy	Blindness (See again)
009	-	+	Anxious	Black Swan (Dressing Room)
011	-	+	Disgust	Anti Christ (Leg Drill)
012	N	-	Sad	Eternal Sunshine of the Spotless Mind (Waking Up)
017	-	+	Anxious	Taken (Boat Fight)
021	-	+	Sad	Artificial Intelligence (Abandoning the Android)
026	+	-	Content	Crash (Invisible Cloak Story)
029	+	-	Amused	Inbetweeners (Nightclub Dance)
031	+	-	Happy	Slumdog Millionaire (I Found You)
034	+	-	Joy	Blindness (Street Rain)
043	N	+	Afraid	The Brave One (Tunnel Attack)
059	+	-	Amused	The Holiday (Phone)
060	N	+	Afraid	The Tree of Life (Dinner Table Fight)

Table 7.2: The Labels are for Valence: Pleasant (+), Unpleasant (-), Neutral (N). For Arousal: Calm (-), Excited (+), Neutral (N)

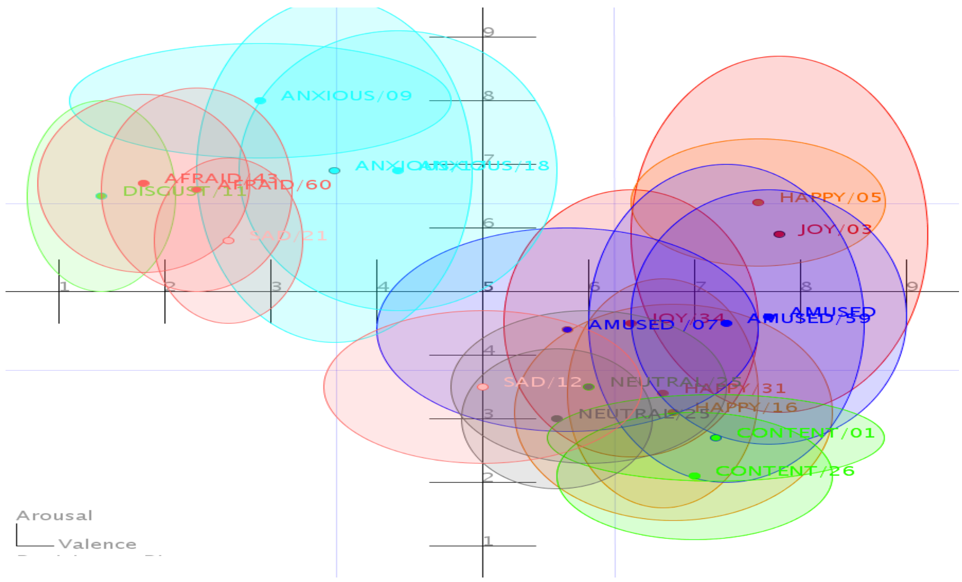


Figure 7.7: Participant (G1) SAM test Valence and Arousal ranges (mean/std) for the 16 selected clips

Figure 7.7 above gathers the across-participant responses for the 16 clips selected for the second stage of the experiment. We have found a mean value from all participants; a dot highlights this value, beside this is the named emotion and also the clip reference number. The standard deviation encircles this on two axes. As Neutral is not being used within this study due to it ambiguous nature, if a variable resides as neutral it will be excluded from the analysis. Above we can clearly denote, that there is an overlap between the certain named emotions, there are no precise single x, y variables or single zone result, and that each felt emotion has more variability for Arousal than for Valence in regards to its locational with the dimensional space.

7.4 Experiment 2: Stage 2

As detailed above participant Group G2, were selected without bias from the initial call out. For the second stage of the experiment this second group (G2), would watch the independently tagged video clips defined by G1, whilst wearing EEG headsets and complete the same SAM test survey in controlled laboratory conditions.

7.4.1 Experiment Procedure.

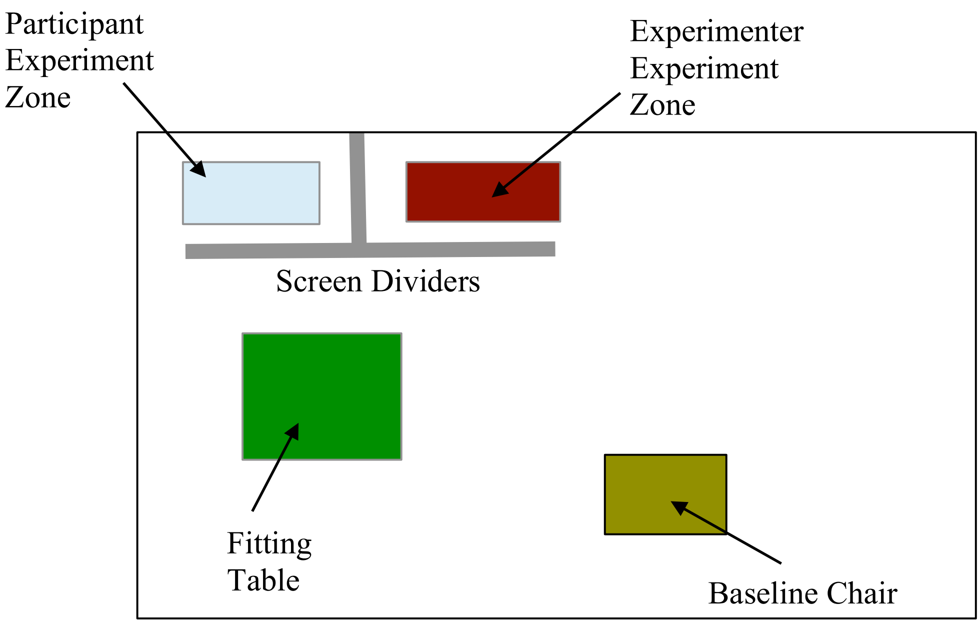


Figure 7.8: the configuration of the experimental space into 3 designated zones

Upon arrival, each participant was led to the experiment space. This was sited in the performance Lab, engineering Building, Queen Mary University of London. The space was sealed off from external noise, interruption, and the lighting was configured to ensure consistent conditions across participants. The space was divided into 3 subsections using screen dividers to meet the experiments needs; a fitting table, a baseline chair, and the experiment zone (see figure 7.8) .

At the fitting table, an information sheet was provided. Here the nature of the experiment and a explanation of the SAM test and its variables were given. Upon signing consent form the participant was fitted with an Emotiv EEG headset. When a clear signal for all of the headsets 14 channels was received, the participant was led to the baseline chair. The baseline chair faced a blank black wall, and the participants were instructed to relax, look ahead, and to minimise movement. No physically restraining or restrictive devices were used, as it was the aim of the experiment to have a natural and relaxed recording. The experimenter then started the 10-minute baseline recording, and sat silently out of view behind a screen divider, providing no distractions. Once the baselines were recorded, the participant was then led to the experiment zone.

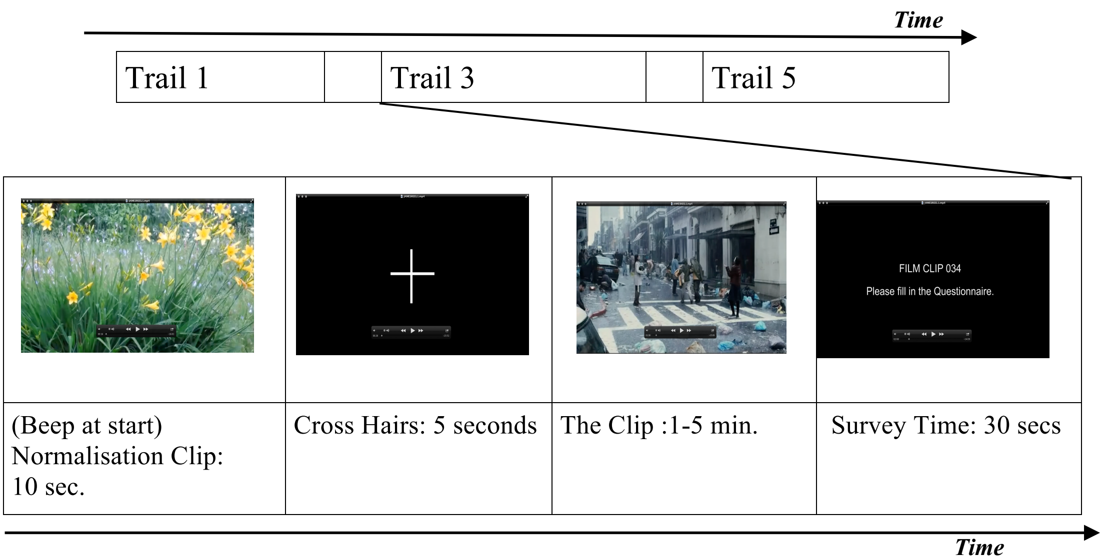


Figure 7.9: The experimental procedure used for all experiments.

Once the participant was seated, the headset was again checked to ensure all electrode contact was intact. The stimulus was show on a 42-inch apple mac screen, connected to a 13 inch MacBook pro laptop from which the experimenter controlled the playback of the content. It was ensured that there was enough space on the table in front of the stimulus monitor for the SAM test to be filled in. 16 SAM tests and writing implements were provided. The researcher who sat opposite, behind a screen divider used a MacBook pro 15-inch to record the live EEG signal in Emotiv's specialized Testbench software. The experimenter applied live markers in Testbench to denote each clips start time, and silently made notes where applicable.

Each trail was formatted as can be seen in the above figure (see Fig 7.9). The experiment operated as such; first a Beep would sound to signify the starting of the trail and to indicate to the participant to watch the screen. A 10 second normalization clip was show in order to remove any bias from the previous stimulus. This was followed by cross hairs for 5 seconds; so the participant may fixate on the central zone of the screen, after which the emotion film clip would be shown. This clip lasted between 1-5 minutes. Finally, text would appear on-screen showing the film clips reference code and instruct the participant to fill in the survey. The time for completing the survey was 30 seconds.

The series of clips were divided into three, approximately 15-minute segments. The rational for this was to allow for any breaks should the participant experience any form of discomfort, and also to allow for any electrode failure to be addressed to prevent loses of large quantities of data.

This procedure was consistent for all participants. The experimenter sat silently behind the screen-divider marking the EEG recording. The participant's welfare was verified between the start of each section of clips. Once completed, the participants were led back to the fitting table, the headset was removed and they were compensated for their time. The whole process including fitting the headset took approximately 2 hours. For the experiments a new set of EEG electrodes and electrodes pads were used. All parts of the headset were cleared between experiments for any oxidation that may have occurred.

7.4.2 Pre-processing & Feature extraction.

The same procedure has been followed throughout this research. Firstly for each participant, the EEG signals for each individual film clip were extracted. For space saving and efficient processing, only the signals for the electrodes of interest F3/F4 were retained and exported to Matlab 2012b.

Here the signals were passed through a bandpass filter to extract only the Alpha Frequency Range (8-13 Hz). This also served to limit the majority of potential artefacts affecting the signals (see section 5). A Sliding Fast Fourier Transform (DFT) was applied at a sample rate of 1024 Hz, with a 50% overlap, giving a frequency representational value of the signal every four seconds.

The Asymmetric Hemispheric Difference algorithm for Valence, and also the prospective Alpha Spectral Power algorithm for Arousal that are both tested throughout this research were applied. (see Table 7.3)

Valence algorithm	log(Alpha, Right Hemisphere) -log(Alpha Left Hemisphere)
Arousal algorithm	(Alpha Right Hemisphere + Alpha Left Hemisphere) /2

Table 7.3: The Algorithms tested for Valence and Arousal detection

Two baseline mean values were created. Firstly, a baseline mean (B) was calculated from the participant's baseline recording. A second baseline mean (ACB) was generated through the concatenation of all EEG clips data for a participant, and a single mean value calculated. This was conducted for each participant for both of the Valence and Arousal vectors.

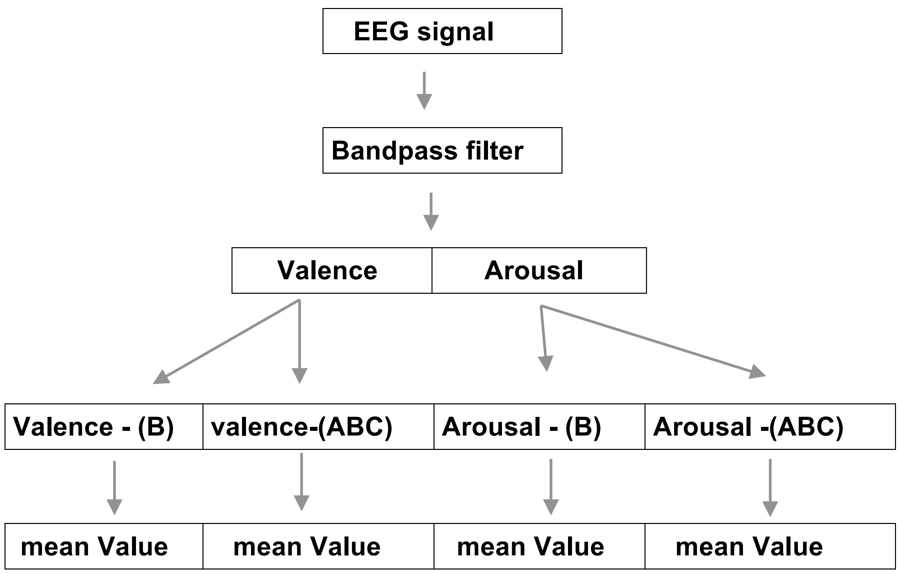


Figure 7.10: A diagram of the method from signal acquisition to output value for classification

For each clip and vector, the baselines (B) and (ACB) were then extracted from each data point in the signal. This served to leave only the variable difference for each clip against baseline (B) and baseline (ACB). A mean value was calculated from the series of values for each clip. Finally this value was reduced to either a positive (+) or negative (-) signifier dependent on whether its value was positive or negative. Naturally, as Alpha is seen as an inverse signal of attention, for Arousal this signifier was inverted. This process was conducted for all participants and respective signals. In turn, the survey results were also reduced to a positive (+) , Neutral (N), or negative(-), signifier dependent on their value ; >5 = (+). ==5 =(N), <5 = (-) (see Fig 7.10). The same procedure has been followed throughout this research.

7.4.3 Survey Results

Two types of data were made available from the experiment, survey data and signal data. From Group 1 (G1) this was in the form of survey data. From Group 2 (G2) both survey and signal data were acquired. In order to make some conclusions, the data was examined in variant ways.

Firstly the survey data across the 2 participant groups was assessed. As we are using only binary conditions on each vector; positive (+) or negative (-), if a mean score for the group ranged <5 the negative signifier (-) was attached, and if >5 a positive signifier (+) was attached. If the value was ==5 then it was considered neutral given a (N) signifier and discarded from the analysis.

7.4.3.1 SAM Test Results

Survey Correlations Between Groups.

	Group 1		Group 2	
Clip	**Valence**	**Arousal**	**Valence**	**Arousal**
001	+	-	+	-
003	+	+	+	-
004	-	+	-	+
005	+	+	+	**N**
009	-	+	-	+
011	-	+	-	+
012	**N**	-	+	-
017	-	+	+	+
021	-	+	-	-
026	+	-	+	-
029	+	-	+	**N**
031	+	-	+	-
034	+	-	+	-
043	-	+	-	+
059	+	-	+	+
060	-	+	-	**N**

Valence Match =93.7%
Arousal Match = 76.9%
Both variable Match = 66.7%

Table 7.4: Survey correlations between groups 1& 2. The Labels are for Valence: Pleasant (+), Unpleasant (-), Neutral (N). For Arousal Calm (-),Excited (+), Neutral (N).

As a first consideration, we can establish whether there is a correlation between the two groups in terms of the returned self-reports by comparing their responses. In

Table 7.4, we can see the results between the mean Valence and Arousal scores for the 2 groups. As indicated above these have been reduced to either a positive or negative value on each vector. For Valence we can note a strong agreement between the 2 groups with a 93.7% correlation. For Arousal this correlation falls to 76%. In both of these instances a random response level is 50%. Thus whilst there is not a complete one to one correspondence between the groups, we do have a good matching rate for Valence, and a reasonable-to-good for Arousal.

7.4.3.2 Emotional Keyword Results.

Keyword Tag Correlations Between Groups

Clip	Group 1 (G1)	Group 2 (G2)	Correlation Y/N
001	Content	Happy	N
003	Joy	Happy	N
004	Anxious	Disgust	N
005	Happy	Happy	Y
009	Anxious	Anxious	Y
011	Disgust	Disgust	Y
012	Sad	Sad	-
017	Anxious	Anxious	Y
021	Sad	Afraid/Sad	N
026	Content	Content	Y
029	Amused	Happy	N
031	Happy	Content	N
034	Joy	Neutral	N
043	Afraid	Disgust	N
059	Amused	Amused	Y
060	Afraid	Disgust/Anxious/ Afraid	N
			6 of 15 = 40 % random is 11.1%

Table 7.5 Tabling of the dominant keyword correlations between the two participant groups.

In a second comparison we can compare the two sets of emotional keywords attached to each film clip. To clarify; following Soleymani et. al (2007), the highest number of emotion keyword registrations for each clip was attached to it. We can note only a one-to-one correlation in 6 of 15 instances (40%), where the random response level is 11.1%.

Contrasting the above two tables (7.4 & 7.5), we can note a clearer correlation rate between the two groups when using the dimensional approach of tagging the clips with Valence and Arousal values, than with the attachment of a keyword. It may be that the wide cultural diaspora of our participant group influences this, as different cultures and nationalities may articulate and associate different names to the same feelings.

7.4.3.3 Survey Correlations

Finally we can also examine the individual scores for the G2 surveys, to consider the level of consensus there is within this group in terms of Valence and Arousal values attached to the selected film clips.

SAM test correlations: Group 2 (G2)				
	Valence		Arousal	
Clip	%	Value	%	Value
001	87.5	+	77.8	-
003	100	+	62.5	+
004	100	-	71.4	+
005	100	+	55.6	+
009	100	-	71.4	+
011	100	-	50	N
012	55.6	+	75	-
017	75	+	75	+
021	87.5	-	50	N
026	100	+	87.5	-
029	100	+	66.7	+
031	100	+	87.5	-
034	71.4	+	87.5	-
043	100	-	87.5	+
059	100	+	57.1	+
060	100	-	60	+

Table 7.6: Percentages of agreed correlations between dominant (+) or (-) values for Group 2 (G2)

In table 7.6 a stronger agreement across participants for reported Valence responses than for Arousal responses is apparent amongst this group. For Valence we have 11 of 16 (68.8%) instances of 100% agreement, whilst there is no 100% agreement for Arousal. There are only 2 clips (012 & 043) where we have a higher matching rate for Arousal over Valence. It should be noted that the random response level is 50%. There are 4 of 16 instances where Arousal correlation between participants is below 59%, inclusive of 2 instances where we reside on the randomness boundary. There is only one instance for Valence where the match rate is lower than 70% and none below 55%. This again contributes to the sense, that Valence may have better classification rates than Arousal, as there seems to be a stronger consensus as to what Valence represents.

7.4.4 Experiment 2: Stage 2 Results

In the following presentation of results, the focus in on the correct classification percentage rates between the EEG signals and SAM test surveys. As mentioned prior, from the EEG data two mean values for each film clip along each dimensional vector of Valence and Arousal have been calculated. Each of these two values is for a different baseline correction method (baseline (B) and baseline (ACB). In turn, these will be evaluated against the survey results in the following formats.

(i) Individual Survey (G2) to Individual Signal (G2)
(ii) Group 1 Survey (G1) to Individual Signal (G2)
(iii) Group 2 Survey (G2) to Individual Signal (G2)
(iv) Group 1 Survey (G1) to Group 2 Signal (G2)
(v) Group 2 Survey (G2) to Group 2 Signal (G2)

Through this series of assessments, we may be able to evaluate the robustness and appropriateness of the algorithms we are using. Successful classification rates under these controlled conditions may give us further confidence in our measures and support our results from experiment 1. Further they may also endorse the use of these methods in further research.

7.4.4.1 Evaluation (i): Individual Survey (G2) to Individual Signal (G2)

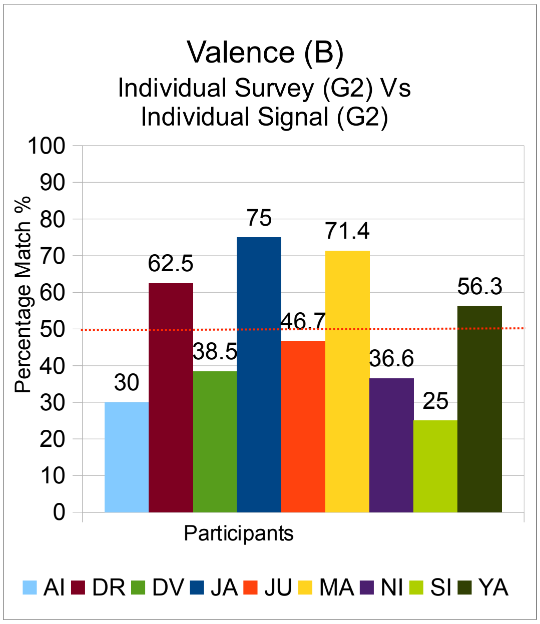


Figure 7.11: Successful classification rates: Valence (B)

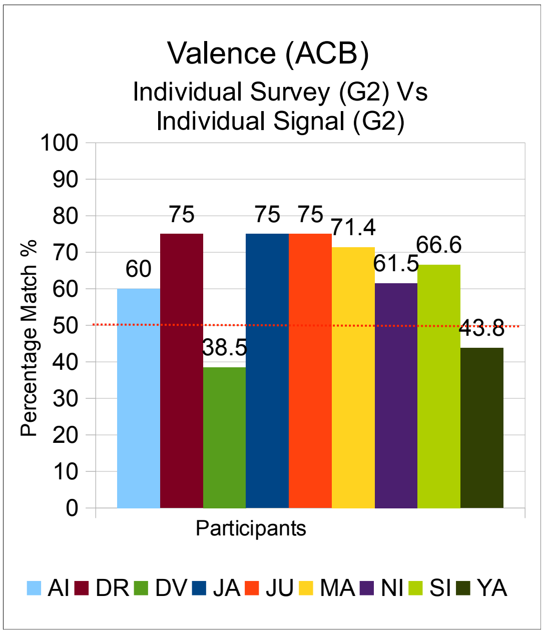


Figure 7.12 Successful classification: rates: Valence (ACB)

Figures 7.11 and 7.12 show the results of the comparison between the Individual Surveys (G2) and Individual Signals (G2) for Valence. Here we can denote a higher correct classification rate when using the (ACB) baseline than for baseline (B). For baseline (ACB), 7 of the 9 participants (77.8%), have a higher than random level of correlation (50%), all of which are either equal to or above 60%. There are 3 participants who share the highest classification rate of 75% (see Fig 7.12).

For baseline (B) (Fig 7.11) we see a significant reduction in correct classification percentages, which are suggestive of chance levels. Only 4 of the 9 participants are above the random rate. Further 1 of these 4 occurrences is marginally above the random level with a rating of 56.3%. The highest individual success score is again 75%, but this is only for 1 individual.

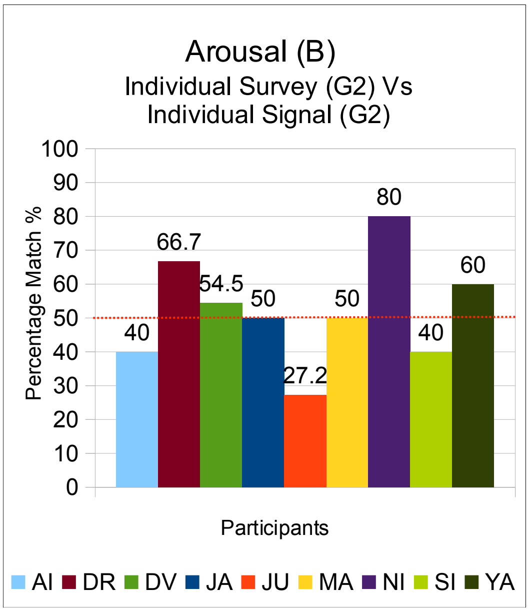


Figure 7.13: Successful classification rates: Arousal (B)

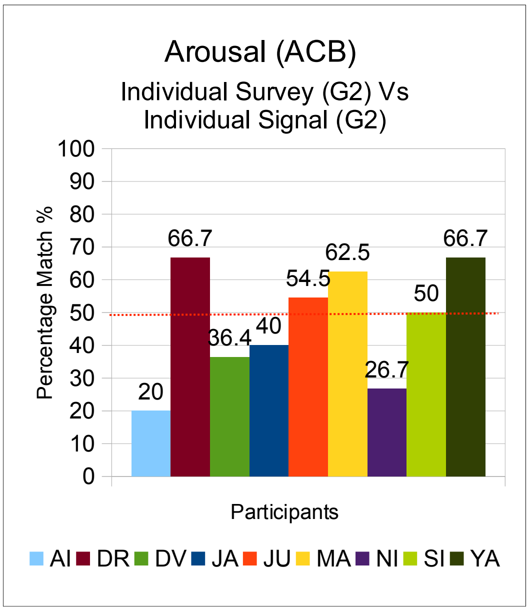


Figure 7.14 Successful classification: rates: Arousal (ACB)

In terms of Arousal classification between an Individual's signal and survey, we find a lower success rate than the rates we found for Valence. For both baseline methods (B), and (ACB) we find only 4 of 9 instances where participants achieve above the random level of 50%. In both cases 1 of these 4 instances are nominally above the random level. For baseline (B) two further participants reside on the randomness boundary, whilst for baseline (ACB) only one. The highest individual success rate is for 1 individual for baseline (B) at 80%. For baseline (ACB), the highest individual rate is shared by 2 individuals at 66.7%.

r-Correlation: Individual Self report (G2) and individual Signal (G2)

Participant	**Valence (B)**		**Valence (ACB)**		**Arousal (B)**		**Arousal (ACB)**	
	R	*P-Value*	*R*	*P-Value*	*R*	*P-Value*	*R*	*P-Value*
AI	**-0.27**	*0.43*	**-0.08**	*0.81*	**0.18**	*0.6*	***0.18***	*0.6*
DR	**0.56**	*0.12*	**0.56**	*0.12*	**-0.47**	*0.2*	***-0.42***	*0.26*
DV	**0.02**	0.99	**-0.03**	*0.99*	**0.03**	*0.93*	***0.03***	*0.93*
JA	**0.54**	*0.03*	**0.52**	*0.04*	**0.15**	*0.57*	***0.15***	*0.57*
JU	**0.53**	*0.04*	**0.58**	*0.02*	**0.14**	*0.6*	***0.05***	*0.86*
MA	**0.46**	*0.07*	**0.41**	*0.12*	**-0.46**	*0.07*	***-0.46***	*0.07*
NI	**-0.15**	*0.66*	**-0.05**	*0.88*	**-0.57**	*0.07*	***-0.57***	*0.07*
SI	**-0.08**	*0.84*	**-0.08**	*0.84*	**-0.46**	*0.18*	***-0.48***	*0.2*
YA	**-0.39**	*0.13*	**-0.3**	*0.26*	**-0.16**	*0.54*	***-0.16***	*0.54*

Table 7.7: Participants Correlation Coefficient r : between individual Signal (abs mean) and Survey for Evaluation (i): Individual Survey (G2) to Individual Signal (G2).

In table 7.7 (above) we consider the linear relationship between the participants signals and SAM test ratings. We created 16 data points for each participant comprised of the mean signal against the SAM value for each clip. We did this for both baselines and both dimensions. We calculated the Correlation Coefficient r, and tabling the results we may surmise that there is not a strong linear one-to-one correlation between them. Of all 9 participants only JU (Valence (ACB), r = 0.58, n = 16, p = 0.02), JA (Valence (B), r = 0.54, n =16, p = 0.03) show a marginal linear relationship for Valence. As a totality for the whole group, this is not so surprising as we have already seen that there is variability between the group in terms of keyword annotation and SAM markers.

This may be due for a number of reasons, the gap between an neural experience and attempts to articulate how one is feeling, or even that each emotion does not have a static reading that can be applied to it.

7.4.4.2 Evaluation (ii) Group 1 Survey (G1) to Individual Signal (G2)

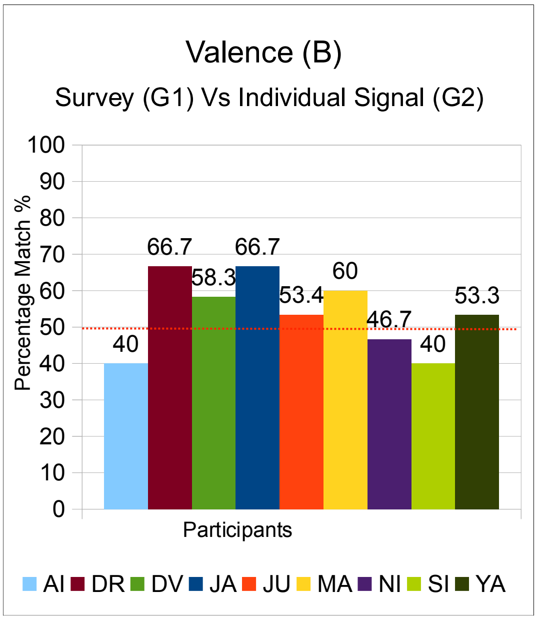


Figure 7.15: Successful classification rates: Valence (B)

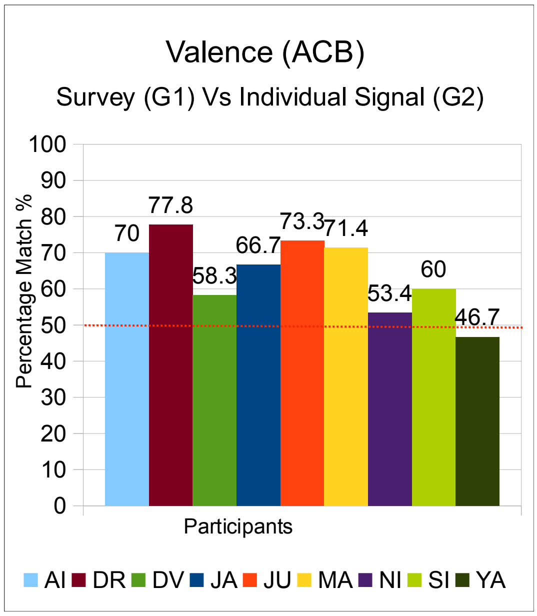


Figure 7.16 Successful classification: rates: Valence (ACB)

In this second comparison we will compare the individual signals of G2 against group G1's mean survey results. Again a more successful classification rate for Valence over Arousal is found. For Valence, using baseline (B) we can view 6 of the 9 participants above the random indication level, although 2 of these operate marginally above this marker (see Fig 7.15). Again the baseline method (ACB) outperforms this, with a significant 8 of 9 participants above the random level with only 1 participant marginally close to this marker with a rate of 53.4%. The remaining participant falls just below the random marker with 46.7% (see Fig 7.16). 5 of G2's participants for baseline (ACB) score above 66%, with the highest individual successful classification rate being 77.8%.

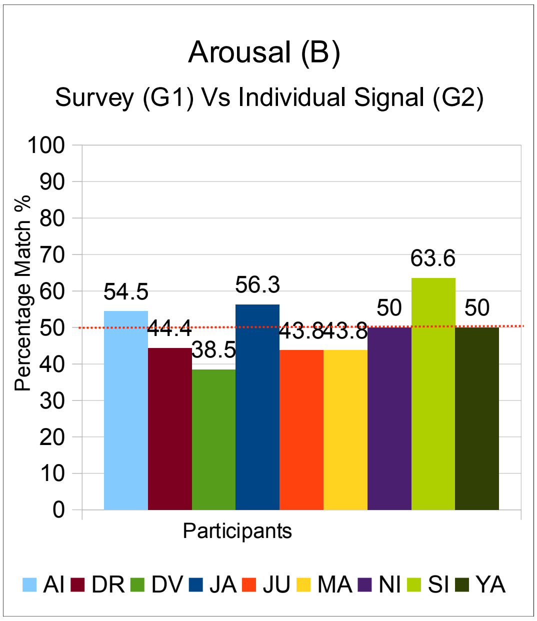


Figure 7.17: Successful classification rates: Arousal (B)

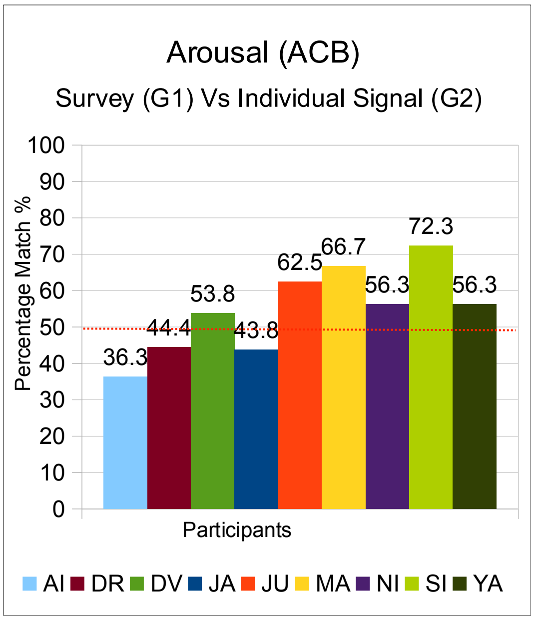


Figure 7.18 Successful classification: rates: Arousal (ACB)

The correct Arousal classification results are again much lower than for Valence. For baseline method (B), only 3 participants are above the random level, of which 2 are close to its vicinity with rates of 54.5% and 56.3%, with a further 2 participants on this boundary (see Fig 7.17). For baseline (ACB), 6 participants clear the random border with, with 3 having rates in excess of 62%, whilst 3 remain close to it location with rates of 53.8%, 56.3% and 56.3% (see Fig 7.18)

Correlation Coefficient r ,Table for group Survey (G1) and Individual Signal (G2)								
	Valence (B)		**Valence (ACB)**		**Arousal (B)**		**Arousal (ACB)**	
Participant	*R*	*P-Value*	*R*	*P-Value*	*R*	*P-Value*	*R*	*P-Value*
AI	**-0.16**	*0.86*	**0.05**	*0.88*	***-0.13***	*0.69*	*-0.13*	***0.69***
DR	**0.5**	*0.17*	**0.5**	0.17	***-0.42***	*0.26*	*-0.18*	***0.64***
DV	**0.19**	*0.54*	**0.19**	*0.54*	***-0.04***	0.91	*-0.04*	***0.91***
JA	**0.35**	*0.18*	**0.33**	*0.2*	***-0.18***	*0.5*	*-0.18*	***0.5***
JU	**0.51**	*0.04*	**0.53**	*0.04*	***-0.11***	*0.68*	*-0.21*	***0.42***
MA	**0.52**	*0.04*	**0.49**	*0.05*	***-0.23***	*0.4*	*-0.23*	***0.4***
NI	**-0.06**	*0.86*	**0.05**	*0.88*	***-0.13***	*0.69*	*-0.13*	***0.69***
SI	**-0.21**	*0.56*	**-0.21**	*0.56*	***-0.07***	*0.84*	*-0.71*	***0.03***
YA	**-0.45**	*0.08*	**-0.36**	*0.17*	***-0.34***	*0.2*	*-0.34*	***0.2***

7.8 : Participants Correlation Coefficient r : between individual Signal (abs mean) and Survey (G1)

In considering the linear relationship between each participants SAM rating and mean signal value for each clip, we found an overall weak correlation. Only one participant (SI) achieved a *r*-value close to a good correlation with -0.71 for the Arousal (ACB) measure (r = -7.1, n =10, p = 0.03) (see table 7.8). This reporting of a weak correlation between the survey and signal data has been consistent throughout our results.

7.4.4.3 Evaluation (iii) Group 2 Survey (G2) to Individual Signal (G2)

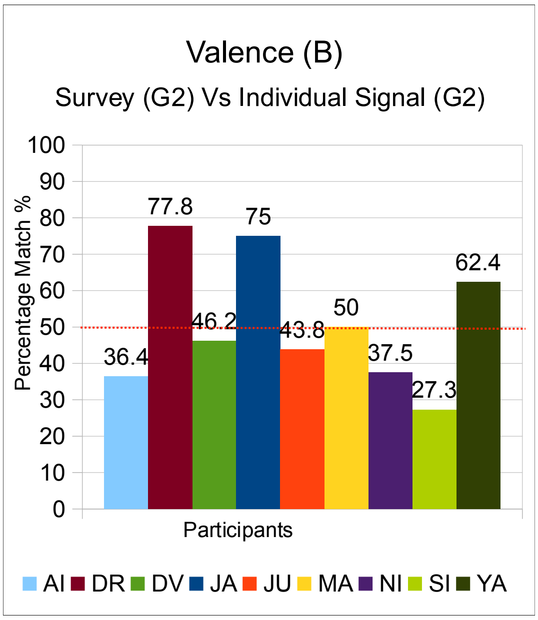


Figure 7.19: Successful classification rates: Valence (B)

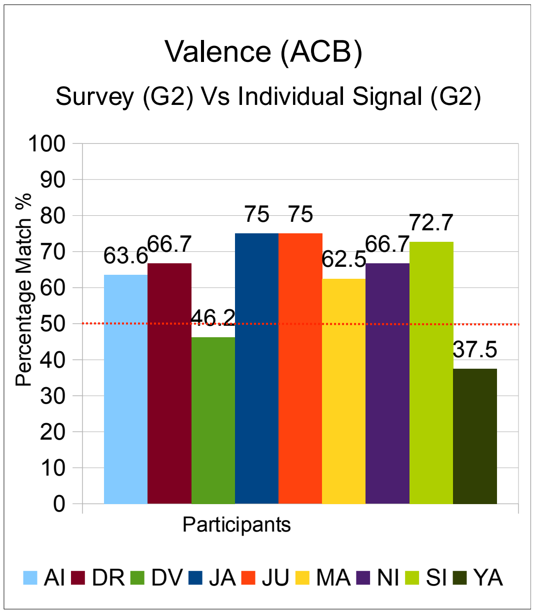


Figure 7.20 Successful classification: rates: Valence (ACB)

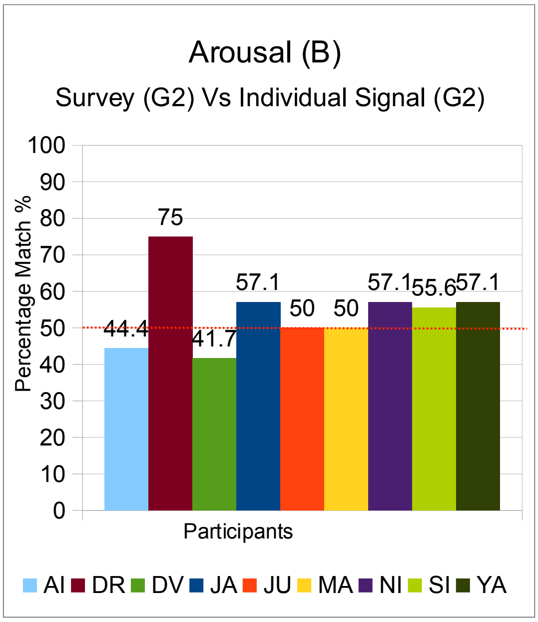


Figure 7.21: Successful classification rates: Arousal (B)

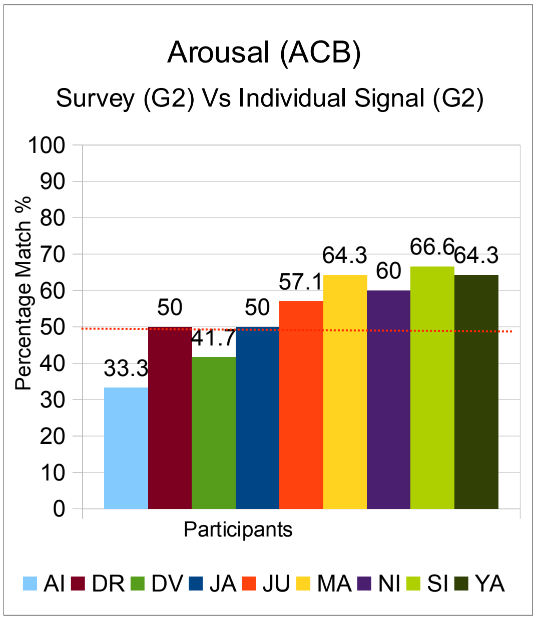


Figure 7.22 Successful classification: rates: Arousal (ACB)

The pattern that emerged from the previous evaluations is repeated when we compare the individual EEG signals from the participants of (G2) with the mean value of surveys from (G2). In terms of Valence using baseline method (B), only 3 of 9 participants score above a random rate although these are all in excess of 62% (see Fig 7.19). This is outperformed by baseline (ACB) where we find 7 of 9 participants with classification rates all in excess of 62% (see Fig 7.20).

For Arousal, with baseline (B), 5 of 9 participants score higher than random rates, with 4 of these edging towards it boundary with scores of 55.6%, 571.%,57.1% and 57.1%. A further 2 participant rest on the 50% boundary (see Fig 7.21). With baseline method (ACB) again 5 of 9 instances are above random, this time with only 1 close to this border with 57.1% and a further 2 participants achieve scores equal to 50%. (see Fig 7.22).

Correlation Coefficient r: for group Survey (G2) & Individual Signal (G2)								
	Valence (B)		**Valence (ACB)**		**Arousal (B)**		**Arousal (ACB)**	
Participant	*R*	*P-Value*	*R*	*P-Value*	*R*	*P-Value*	*R*	*P-Value*
AI	**-0.16**	*0.63*	**-0.04**	*0.9*	***-0.19***	*0.58*	**-0.16**	*0.67*
DR	**0.72**	*0.03*	**0.72**	*0.03*	***-0.68***	*0.05*	**-0.32**	*0.41*
DV	**0.17**	*0.58*	**0.17**	*0.59*	***0.06***	*0.85*	**0.08**	*0.8*
JA	**0.38**	*0.15*	**0.36**	*0.17*	***-0.16***	*0.56*	**-0.16**	*0.56*
JU	**0.47**	*0.06*	**0.49**	*0.05*	***-0.11***	*0.7*	**-0.22**	*0.42*
MA	**0.53**	*0.04*	**0.49**	0.06	***-0.3***	*0.26*	**-0.3**	*0.26*
NI	**-0.16**	*0.63*	**-0.04**	*0.9*	***-0.19***	*0.58*	**-0.19**	*0.58*
SI	**-0.06**	*0.86*	**-0.06**	*0.86*	***-0.68***	*0.03*	**-0.76**	*0.02*
YA	**-0.46**	*0.07*	**-0.51**	*0.04*	***-0.31***	*0.25*	**-0.26**	*0.34*

Table 7.9 Participants Correlation Coefficient r : between individual Signal (abs mean) and group survey (G2).

As with the previous 2 evaluations, an over all weak value for correlation coefficient r, can be found across participants (see table 7.9). Only participants; DR: (Valence (B) & (ACB) r = 0.72, n = 9 , p= 0.03), and (Arousal (B), r = -68 , n = 9, p = 0.05) , and SI, (Arousal (B), r = -0.68, n =10, p = 0.03) and (Arousal (ACB), r = -0.76, n = 10, p = 0.02) demonstrate forms of a liner relationships between their surveys and signals.

7.4.4.4 Evaluation (iv) Group 1 Survey (G1) to Group 2 Signal (G2)

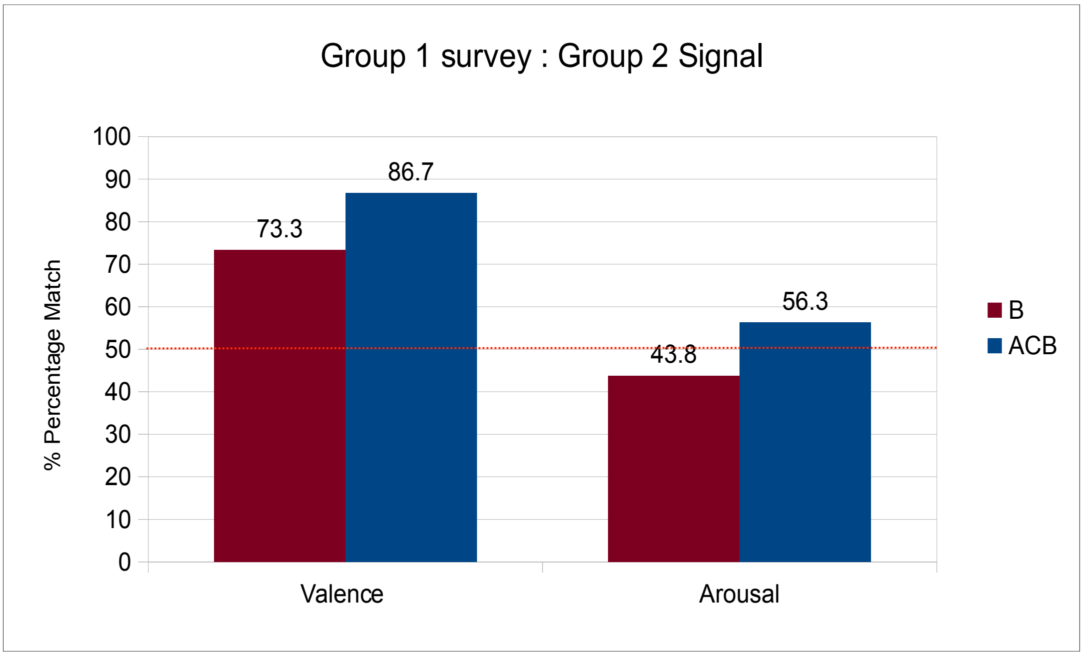


Figure 7.23: Successful group classification rate: Group 1 (G1) survey Vs. Group 2 (G2) signal

Having seen the rates of successful classification when assessing group 2's (G2) signal data as individuals, we may also average the group's population signals and treat them as a single entity. Here, when we treat the individuals as a group by producing a mean value for signals and surveys across the whole group, we can begin to see a clear differentiation between the classification rates of correct Valence and Arousal detection. In the instances of both baselines methods (B) and (ACB), we have good binary classification rates of, 73.3% for (B), and 86.7% for (ACB). For Arousal classification, we find poorer results, with figures 43.8% and 56.3% that are nominally around the level of randomness.

7.4.4.5 Evaluation (v) Group 2 Survey (G2) to Group 2 Signal (G2)

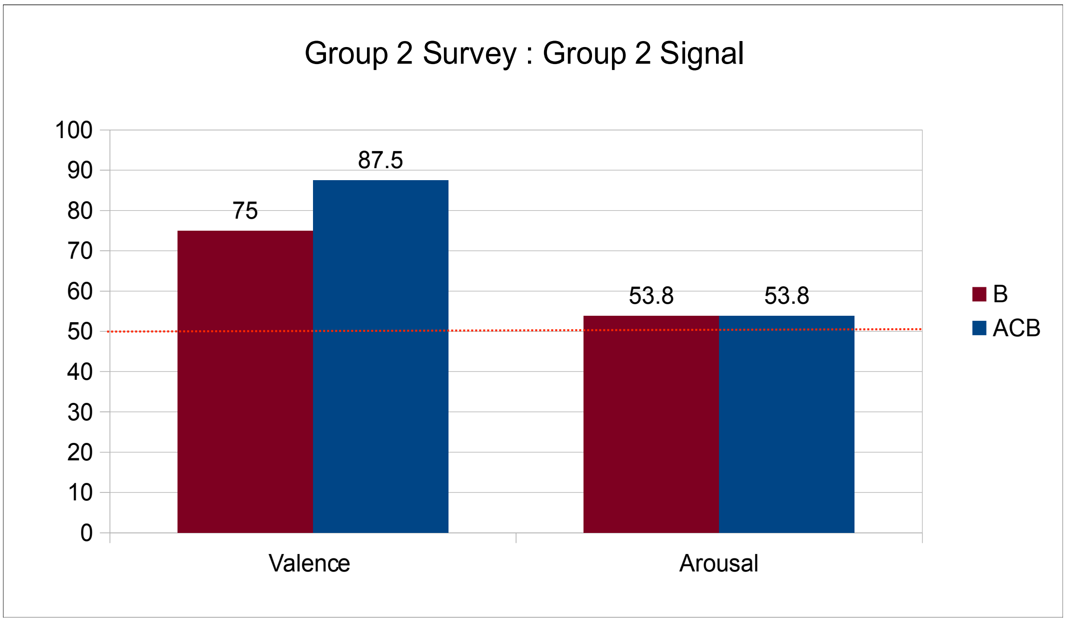


Figure 7.24: Successful group classification rate: Group 2 (G2) survey against Group 2 (G2) signal

Echoing the comparative findings of evaluation (vi) we have similar yet slightly improved results when classifying G2 (as a group) signal against G2 (as a group) survey. For baseline (B) we have a Valence classification score of 75 %, whilst for baseline (ACB) we have the highest evaluated correlation of 87.5 %. Here, the correct derived Arousal classification rate is 53.8% for both baselines, which is only marginally above the random level. In both of these group analyses, we can clearly see better results for Valence detection over Arousal detection.

Thus, when we consider treating (G2), as a single entity we can note a similar pattern to when we treat G2 as individuals. There is a marginal-to-good correlation between signal and survey for Valence, and in particular with baseline (ACB). Simultaneously for Arousal using this format of experiment we find less confident results, and the results operate at chance levels. It may be questioned whether the conditions of the experiment may have an impact on the levels of emotion investment, and in turn the Arousal results, this is an aspect which would require further detailed investigation.

r-Correlation Summary : Group Surveys and Group Signals								
	Valence (B)		Valence (ACB)		Arousal (B)		Arousal (ACB)	
	R	P-value	R	P-value	R	P-value	R	P-value
Group 1 Survey: Group 2 Signal	**0.39**	*0.14*	**0.54**	*0.03*	**-0.26**	*0.34*	**-0.22**	*0.4*
Group 2 Survey: Group 2 Signal	**0.37**	*0.16*	**0.49**	*0.06*	**-0.26**	*0.33*	**-0.36**	*0.17*

Table 7.10: Summary Table of group Correlation Coefficient r : between Group Signal (abs mean) and Group survey (means) (G1) and (G2)

Finally, as with all other evaluations questioning a linear relationship between signals and surveys, we again find a weak correlation between them (see Table 7.10). This supports the notion that there is not a strict one-to-one linear relationship, where specific annotated variables are consistent with a discreet value in the signal. As we have seen above in all forms of this analysis we cannot consolidate such a relationship.

7.4.4.6 Results Summary.

As outlined above in the results, for Valence we have consistently scored marginal-to-good successful classification rates across a number of different formats of evaluation. This is regardless of whether the participant populations were treated as individuals or as a group. For the evaluation (i), where participants were treated as individuals in both signal and survey, using baseline (ACB) we found that 7 of the group of 9 participants scored above 59% successful classification rates, with 3 individuals scoring 75%, and another 71.4%.

Again for baseline (ACB), in evaluation (ii), where the surveys were represented by the averaging of (G1) SAM test responses, 8 of 9 participants scored above a 53% success rate, with 6 of these scoring in excess of 59% .The remaining 1 participant had results just below random with 46.7%. For the same baseline (ACB) in evaluation (iii) where individual signals (G2), were compared to an averaging of SAM test scores (G2) we found 7 of the group had success rates above 62%. Evaluation (iii) may be seen as the most consistent method amongst individual participants for higher success rates followed closely by evaluation (i). Thus, we may state through our results that Valence (ACB) consistently returns marginal-to-good successful classification rates for individual signals against survey, and provides some sense of confidence that this measure has some validity for gauging this form of response via EEG, and with the headset we are using. These confidence levels may be tested for further increase by using a larger participant group, and also with further consideration given to creating more distinct polar Valenced (pleasant -unpleasant) stimuli.

With Valence levels for baseline (B), we found a drop in these success rates. The best classification rates were achieved with evaluation (ii): individual Signals against G1 surveys. Here, 6 of 9 participants achieved success rates higher than 53%. For evaluation (i) only 4 participants achieved rates in excess of 56.3%, and for evaluation (iii) only 3 participants achieved above random scores. This may have been due to the larger difference in conditions between the baseline recording and watching an onscreen stimulus.

However, the Valence results for both baselines where the evaluations treated the populations as group signals against both of the groups surveys, demonstrate a good classification result of over 73% in both baseline instances (see table 7.23).

For Arousal, the success rates were much lower than for Valence. However for baseline (B) evaluation (iii) of individual signals (G2) against the averaging of the G2 surveys, resulted in 5 participants exceeding a 55% classification rate with a further 2 on the 50% boundary. This was the best performer, followed by the evaluation (i); individual signals to individual survey, with 4 participants achieving success rates in excess of 54% and a further 2 participants on the 50% random level.

These success rates were improved for baseline (ACB). Using evaluation (ii) where individuals signals were compared to the averaged SAM test score of group (G1), 6 of 9 participants rates were in excess of 53%. This best performer was followed by evaluation (iii) this time comparing individual signals against group (G2) averaged SAM test scores, where 5 participants scored above 57% with a further 2 participants on the 50% boundary.

Averaged Results for Individual Signals in Evaluations (i) - (iii)								
	Valence (B)		**Valence (ACB)**		**Arousal (B)**		**Arousal (ACB)**	
Evaluation	mean	*std*	mean	*std*	mean	*std*	mean	*std*
(i)	49.11	*18.10*	63.98	*13.67*	52.04	*15.74*	47.06	*17.26*
(ii)	53.90	*10.14*	64.09	*10.08*	49.43	*7.77*	54.71	*11.64*
(iii)	50.72	*17.53*	62.88	*12.95*	54.22	*9.67*	54.14	*11.34*

Table 7.11: Averaged participants classification %, success for evaluations (i) - (iii), where individual's signals were assessed.

For simplicity of viewing of we may average the success rates for each evaluation across the group, to clearly see the difference between the two baseline measures. As table 7.11 shows, there is a greater consistency for Valence, baseline (ACB) across the evaluations, and baseline (ACB) for Arousal. This is different than the results for experiment 1 (see Chapter 6) where baseline (B) was the best performer for Arousal. From all the evaluations, where individual signals were used (i)-(iii) we can see the baseline (ACB) in evaluation (ii) group 1 (G1) surveys to individual signals (G2) was the best method performed.

Number of Participants with Higher than Random Classification Rates in Evaluations (i) - (iii)				
	Valence (B)	**Valence (ACB)**	**Arousal (B)**	**Arousal (ACB)**
Evaluation	*>51%*	*>51%*	*>51%*	*>51%*
(i)	*4 of 9*	*7 of 9*	*4 of 9*	*4 of 9*
(ii)	6 *of 9*	8 *of 9*	3 *of 9*	6 *of 9*
(iii)	3 *of 9*	*7 of 9*	5 *of 9*	5 *of 9*

Table 7.12: The number of individual participants achieving higher then random successful classification rates in evaluations (i) - (iii).

Table 7.12, shows the number of individual participants achieving higher than random classification rates in each evaluation. In evaluation (ii) for baseline (ACB) 88% of participants had higher rates than 51% for valence, and 66.7% had higher rates than 51 % for Arousal.

Summary of Successful Classification rates for Group survey and Group signals.				
	Valence B	**Valence ACB**	**Arousal B**	**Arousal ACB**
Evaluation	%	%	%	%
(iv)	*73.3*	*86.7*	*43.8*	*56.3*
(v)	*75*	*87.5*	*53.8*	*53.8*

Table 7.13. A summary table of successful classification results, for group comparisons, evaluation (iv) Group 1 (G1) surveys against Group2 (G2) signals, and evaluation (v) Group 2 (G2) surveys against Group2 (G2).

As a final table, When the group (G2) signals were averaged and compared against the averaged group surveys, in both instances of Valence and Arousal baseline (ACB) outperformed or matched baseline (B) for successful classification. The best performance achieved was evaluation (v) for Valence 87.5% and evaluation (iv) 56.3% for Arousal (see Table 7.13).

7.5 Conclusion and Discussion

Here we conducted an experiment comprised of two stages. In the first stage a group of 10 participants watched 38 film clips and provided SAM test ratings and a keyword tag for each. From these responses 16 films clips were selected based on predominant keywords, with 2 clips representing each of the eight emotions; Afraid, Amused, Anxious, Content, Disgust, Happy, Joy, and Sad.

A second group of 9 participants, wearing EEG headsets, in controlled conditions watched these objectively selected clips, and gave their own set of survey data. Evaluations were then conducted between signal and survey in a number of evaluations. The key Questions asked were.

(i) Valence, is the AHD method for detection reliable?

(ii) **Arousal** is the ASP method for detection reliable? is there any basis for neural power levels in the alpha region being indicative of Arousal?

(ii) Will having a larger experimental **population** allow for successful classification rates for both individuals and as a group?

(iv) Is the **technology** robust and reliable for this form of experiment?

(v) Which **baseline correction** method is most appropriate for this setting?

In laboratory conditions we were able to confirm some degree of reliability of Valence detection via our methods. This variable was consistent in achieving marginal-to-good classification rates for baseline (ACB) across all individual signal evaluations, where a minimum of 7 of the 9 participants had successful classification rates higher than 51%. As a group this measure was always above 87%. This consistent performance is in line with peer publications and results. Thus we can state a level of confidence in the use of this measure, and it may be by expanding our use of electrodes to further electrode pairings over the frontal lobe that we may be able to have increased success.

For Arousal our findings are not so clear. Whilst for one evaluation (evaluation ii) with baseline (ACB) 6 of the 9 participants had successful classification results higher than 53%, this level is not achieved consistently across the various evaluations in the same way as it did for Valence, thus it does provide some uncertainty, and Arousal may be considered as operating close to chance levels in this situation.

As has been expressed throughout this thesis, Arousal detection seems more problematic than for Valence detection, this may arise from a number of reasons associated with this dimension that have been outlined. In the context of this particular experiment we may further consider the causal impact of our laboratory setting on our data. Film and Cinema have structures and devices that intend to increasingly involve a viewer in the representational reality they are depicting. A short film clip extracted from its larger context and structural function of additive involvement may prevent a participant from building any from of relationship or emotional investment to the characters and the action. When this is combined with the stop - start nature of the experiment, it may generate a situation whereby the participant is unable or unwilling to become sufficiently emotionally invested or aroused by the stimulus. The impact of being reminded that one is under experimental conditions may further be preventative of the subtleties of Arousal levels registering which would affect our results.

In future experiments of the same method, it may be necessary to include enhanced polar extremities of stimulus to gain clearer insights. Alternatively, it may be more appropriate to show a whole film and then take annotations by replaying audio-clips from the film, as was the procedure for experiment 1(see chapter 6). The genre of short film, where productions can last anything from 1-30 minutes may also offer another possible solution for consideration.

The reports of Arousal as a more personally contextualised vector is also reflected in the differences between participants self-reports, and this may lead to lower classification rates when evaluations are considered as group. Thus whilst in this setting some success has been achieved for Valence using the AHD model, where we found a marginal-to-good results, for Arousal we found only indeterminacy with results close to random. Thus it is felt that the nature of the relationship between the level of neural activity recorded as a montage of electrodes F3/F4 and any resultant relationship represented as Arousal levels needs further inquiry, and perhaps in both settings.

We again questioned the robustness and reliability of the chosen technology, and again this was affirmative, the only issue arising from having to re-moisten the felt pads of the electrodes during experiments. It also should be noted that we are using the minimal amount of electrodes possible (F3/F4) and it may be by using more of the headsets electrode pairings such as Fp1/Fp2, and F6/F7 that are also positioned over the frontal area might lead to enhanced results.

Our final question re-visited the issue of the correct baseline procedure (B) or (ACB). We found that (ACB) out performed (B) in all instances. This may be due to the difference between the baseline state and the in-stimulus state. It may be that watching a screen (which flickers at specific rates) may have some form of brain entrainment through sympathetic resonance, which would always affect the results. It is felt by the researcher that this in-determined aspect may require further detailed study. Whilst baseline (B) could be enhanced by taking a third baseline after the stimulus for comparison to attempt to counter any potential brain entrainment from such a stimulus, it is considered that a default position should be to consider taking all 3 baselines, before, during, and after the stimulus. In future research it may be important to further carefully consider the difference between baseline recording conditions and the stimulus in the experimental design.

Thus from this experiment we may state some validity of being able to determine emotional Valence indicators using the AHD method in laboratory conditions, yet with little confidence in the consistency of Arousal indicators using ASP values.

CHAPTER 8

Conclusion

8.1 Conclusion.

Over the prior seven chapters we have detailed the progression of a broad idea into a formal research project. Our intentions were to assess whether it was possible to neurally detect emotional responses to creative and cultural art forms via EEG. This was with a view to attempt to find some form of objectivity to these highly subjective experiences, which may allow some degree of transparency into these aesthetic responses. We aimed to define a process for this. Thus within this context of assessing whether low-cost, portable EEG devices can impact our understanding of cultural experiences in the wild?, we found the following.

We found our results to be consistent with peer literature. It seems that only Valence can be reliably measured, and only with a good degree of confidence in laboratory based studies. When moving into complex settings with complex stimulus we increase the number of unknowns that may disguise our true signal. This is reflected in our comparative experiments where we found less successful Valance results for the experiments in natural settings than for the controlled laboratory settings.

Given that even simple dimensional models of emotion depend on being able to measure both Valence and Arousal, this undermines the value of EEG techniques for assessing emotional responses to Artworks as a standalone technology. It may be more profitable to incorporate EEG within a multi-modal set up where it may contribute to a wide range of measured readings. Whilst this would operate best in controlled conditions, adding further sensors may potentially be able to provide the necessary support for moving beyond the laboratory.

We have also shown through our experiments that in controlled circumstances people do agree quite highly on self-reported measures of Valence and Arousal so this seems like a better singular method than the single modality of EEG for assessing these dimensions. Despite this, our first studies and discussion of responses to a range of artworks highlight the limitations of these rather simplistic models of emotional response for doing this kind of analysis.

Also the later studies show that even where people agree on the Valence and

Arousal dimensions they still assign different descriptive labels to stimulus clips. So, to make significant progress in this area we rather need techniques that can make much more subtle differentiations of response. This would require much more sophisticated techniques that are truly capable of use in the normal circumstances under which people typically encounter artistic works.

Our second interlinked intention was to consider whether through our EEG detection methods we might be able to make the emotions available as a form of creative material for practitioners. In the context of making, having access to the singular dimension of Valence is highly valuable. As noted in our Introduction (see chapter 1.2), artists have made highly innovative and exciting works using simple interpretations of the EEG signal, such as Alpha Spectral Power to drive aesthetic works. Thus in a similar way the Valence level indicator provides the starting point to begin making innovative works in creative contexts.

Throughout the latter stages of this research project the obtained EEG signals and Valence and Arousal classifications were explored through many creative works to consider such usability and functionality.

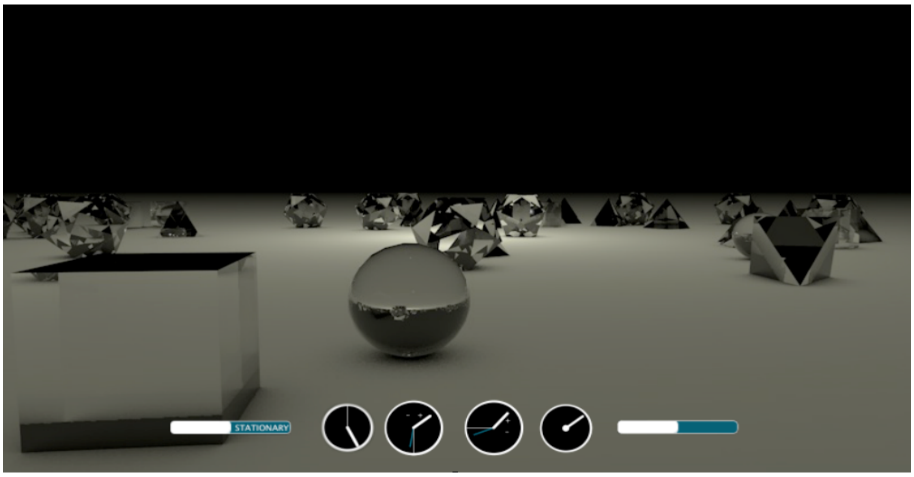


Figure 8.1: A still image from Relentless 2014. (11 mins 24 secs) ; a creative visualisation of neural emotional response signals to James Turrell's light installation Kermandec. 2014, recorded at the Pace Gallery. London.

Figure 8.1 shows just one example of such a work created from the pilot study recording to James Turrells Kermandec exhibited at the Pace Gallery (see section 4.6.7). Valence and Arousal classifications were extracted at four-second intervals and then composited in a 2-dimensional space. A primitive shape represented each of the combined classifications. ; Sphere = class 1 (V+, A+), Pyramid = class 2 (V+, A-), Octagon = class 3 (V-, A-), Cube = class 4 (V-, A+). Every four seconds a new classification shape drops into the viewing frame to disrupt the scene. The work is intentionally relentlessly repetitive, and tries to metaphorically convey how neural 'instructions' may be continuous ‘dripping’ through the body to shape our worlds. The fallen classes remain in view representing the storage of emotional memories for future access. This restrained example points to the endless and exciting possibilities of work that may be generated from such data.

8.2 Future Research.

There are a number of possible avenues for further EEG research whilst working within the limitations of the dimensional framework; (i) Increasing the number of electrodes in the frontal lobe region, (ii) Exploring the further frequency bands of Beta, Theta, Delta and Gamma, (iii) Incorporation of EEG within a multimodal set up. All of these may serve to increase classification accuracy and success rates for Valence and allow for the exploration of Arousal signatures.

In terms of experimental settings, by focusing on one aspect of cultural production, such as short Moving Image works, experiments could be configured in both controlled and natural settings with on par repeatability. This could also incorporate the development of online real time automatic classification system. Finally the further excavation of theoretical frameworks may allow an advancement beyond the dimensional model.

Appendix A

Experiment 1: Paperwork.

A.1 Information Sheet.

Information sheet

Research study [title]: Generating Affective 'Avant-Garde Animation' through Neurological and Physiological interaction relationships

We would like to invite you without any obligation to be part of the research project described below. You should only agree to take part if you would like to. No disadvantages will arise should you decline to be involved.

Before making your decision, please read the following information carefully. This will inform you of the reasons for the research and inform you of what you will be asked to do if you decide to take part. Please do not hesitate to ask if there is anything that is not clear to you, or if you would like further information.

If you decide to take part, you will be asked to sign the attached form as proof of your agreement. Please note that you are still free to withdraw from the research study at any time and without giving a reason.

Details of study

This Study is designed to investigate the neural correlations of dynamic emotional states through Electroencephalographic (EEG) and physiological (q-sensor) readings, in order to determine control measures for driving Affective Interactive Animations.

The study will take place at a central London Theatre. Upon entering the Theatre, you will be guided to an quiet area, where you will remain for up to a maximum of 10 minutes, in order to relax and be clear minded. The experimenter will then fit you with an EEG wireless headset, to your head, which measures real-time electrical activity in the brain, a Biosensor to your wrist, which measures skin conductivity levels, and a small microphone in order to record the ambient sound of the performance. The experimenter will take a short baseline recording and ensure that all the sensors are working correctly. Once complete you will be guided to your seat in the Theatre auditorium.

The experiment requires you to relax and simply enjoy the live theatrical performance. The performance will last for approximately 1.5 hours. Upon completion, you will be guided to a quiet space where you will be asked to recall some of your thoughts, emotions and feelings in response to the performance. A script of the performance may be provided to assist you in this. Finally a short questionnaire will be presented which will conclude the experiment. The experiment will last approximately 3 hours. If you do decide to take part, you will be given this information sheet to keep and be asked to sign a consent form.

A.2 Consent Form.

Queen Mary
University of London

Consent form

Please complete this form after you have read the Information Sheet and/or listened
to an explanation about the research.

Title of Study: Generating Affective 'Avant-Garde Animation' through Neurological and Physiological interaction relationships

Queen Mary Research Ethics Committee Ref: ________________

• Thank you for considering taking part in this research. The person organizing the research must explain the project to you before you agree to take part.

• If you have any questions arising from the Information Sheet or explanation already given to you, please ask the researcher before you decide whether to join in. You will be given a copy of this Consent Form to keep and refer to at any time.

• I understand that if I decide at any other time during the research that I no longer wish to participate in this project, I can notify the researchers involved and be withdrawn from it immediately.

• I consent to the processing of my personal information for the purposes of this research study. I understand that such information will be treated as strictly confidential and handled in accordance with the provisions of the Data Protection Act 1998.

• I consent to the use of images taken during the experiment to be used only in publications relating to this research (optional please tick the appropriate box, below)

☐ I agree to images of my interaction being used in publications related to this study.

☐ I do not agree to images of my interaction being used in publications related to this study.

Participant's Statement:

I ____ agree that the research project named above has been explained to me to my satisfaction and I agree to take part in the study. I have read both the notes written above and the Information Sheet about the project, and understand what the research study involves.

Signed: Date:

Investigator's Statement:

I ____ confirm that I have carefully explained the nature, demands and any foreseeable risks (where applicable) of the proposed research to the volunteer

A.3 Pre-Performance Questionnaire

Queen Mary
University of London

PRE PERFORMANCE Questionnaire.
Josephine & I , The Bush Theatre. London

AGE [ ]

GENDER MALE / FEMALE * please circle

CURRENT OCCUPATION [ ]

HANDEDNESS LEFT | RIGHT | AMBIDEXTROUS * please circle

HOW WOULD YOU DESCRIBE YOU ETHNIC BACKGROUND ? * please circle.

(a) Black or Black British
Caribbean
African
Any other Black background.
background

b) White
British
Irish
Any other White

(c) Asian or Asian British
Indian
Caribbean
Pakistani
African
Bangladeshi

Any other Asian background within

(d) Mixed
White and Black

White and Black

White and Asian

Black and Asian
Any other Mixed

(e) Chinese or other Ethnic Group
Chinese
Any other Ethnic Group

(f) Rather Not Say
Rather not say.

Queen Mary
University of London

DO YOU REGULARLY ATTEND ANY OF THE FOLLOWING CULTURAL EVENTS ?
*please tick the appropriate boxes.

		NEVER	RARELY	OCCASIONALLY	OFTEN
ART GALLERIES	Y/N				
ART FESTIVALS	Y/N				
MUSIC PERFORMANCES	Y/N				
MUSIC FESTIVALS	Y/N				
MUSEUMS	Y/N				
COMEDY	Y/N				
CINEMA	Y/N				
DANCE PERFORMANCES	Y/N				
THEATRE PERFORMANCES	Y/N				

HAVE YOU ATTENDED ANY THEATRE PERFORMANCES IN THE LAST YEAR Y | N

HOW MANY?

WHAT WAS THE LAST THEATRE PERFORMANCE YOU SAW + YEAR (* enter below)

HOW LONG DID THE PERFORMANCE STAY WITH YOU (*approx)

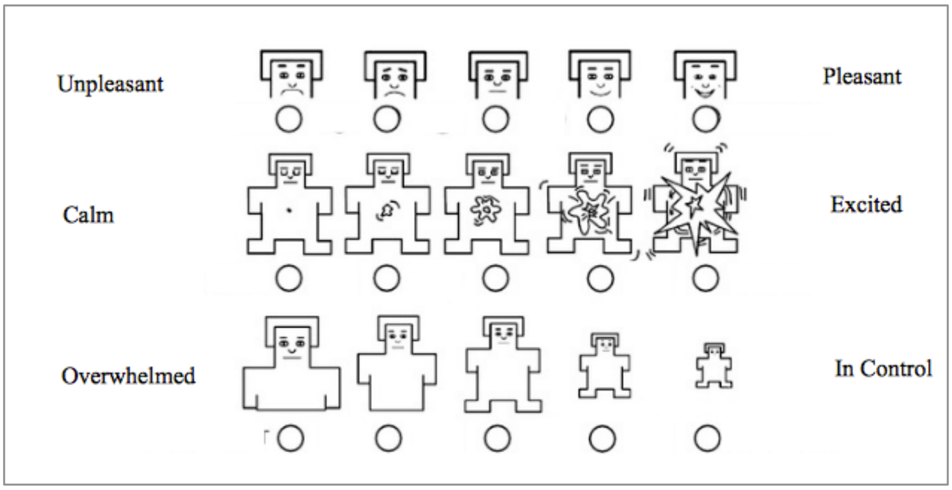
A.4 SAM Test

Sound clip :

Length :

Descr :

Section A:



Section B: Was the feeling consistent throughout the clip. (* please briefly describe)

Section C: Are there any words (emotional expression) you could use to describe how you felt.

Section D : Any Other comments.

A.5 Post Experiment Questionnaire

Queen Mary
University of London

Post Experiment Questionnaire

Section A : Headset

(i) Was the fitting of the EEG headset easy ?

(ii) Was the EEG headset comfortable to wear ?

(iii) Were you always aware that you were wearing the headset ?

(iv) Did the headset distract from the performance at any time ?

(v) Did the headset become uncomfortable at any point ?

(vi) Did anyone notice / comment / ask questions about you wearing the headset ?

(vii) Did wearing the headset make you feel self-conscious at all ?

Queen Mary
University of London

Section B : Sound Clips

(i) Were the replayed sound clips clearly audible

(ii) Upon Hearing them, were you able to remember the point in the performance they occurred

(iii) Did the sound clips ....

(a) bring back the emotions you felt when you watched the live performance
(b) give new emotions
(c) other (* please briefly explain)

(iv) Did you feel the sound clips were .....

(a) too short for the SAM test
(b) just right for the SAM test
(c) too long for the SAM test

(v) did you feel that there were ;

(a) too few sound clips
(b) about the right amount of sound clips
(c) too many sound clips

Queen Mary
University of London

Section C : questionnaire

(i) was the explanation of the SAM TEST clear and easy to understand ?

(ii) did you find it easier to fill in the SAM test or to write the emotion in words ?

(iii) would you have preferred to have been informed of the SAM method before the show ?

(iv) if you answered yes to the above question (iii), do you think this would have made a difference to how you filled out the SAM test ?

(v) Did you find any of the questions on the test difficult ? (* please explain)

(vi) Did you find any of the questions on the test too open ended ? (* please explain)

Appendix B

Experiment 2: Paperwork

B.1 Stage 1 Information Sheet.

Queen Mary
University of London

Information sheet

Research study [title]: Generating Affective ‘Avant-Garde Animation’ through Neurological and Physiological interaction relationships (**QMREC2012/45**)

We would like to invite you without any obligation to be part of the research project described below. You should only agree to take part if you would like to. No disadvantages will arise should you decline to be involved.

Before making your decision, please read the following information carefully. This will inform you of the reasons for the research and inform you of what you will be asked to do if you decide to take part. Please do not hesitate to ask if there is anything that is not clear to you, or if you would like further information.

If you decide to take part, you will be asked to sign the attached form as proof of your agreement. Please note that you are still free to withdraw from the research study at any time and without giving a reason.

Details of study

The larger aims of this Research are to investigate the neural correlations of dynamic emotional states through Electroencephalographic (EEG) readings, in order to determine control measures for driving Affective Interactive Animations.

This particular experiment involves the tagging of Film and Movies clips with emotional labels. For the experiment you will be asked to watch a number of clips and fill in a provided questionnaire for each. The experiment can be conducted on a designated computer in G2, MAT Computer Lab, Engineering Building , Queen Mary, University of London on either Friday 14th February and Wednesday 19th February between 10am - 4pm..

Alternatively the experiment can be provided on a DVD with experiment instructions which can collect on Wednesday 12th February or Friday 14th, to do be done in an quiet environment and time chosen by yourself. The experiment should last no longer than 2.5 hours, and can be conducted in one session or shorter sessions as desire. It is recommended that no shorter session be less than 20 mins. If you do decide to take part, you will be given this information sheet to keep and be asked to sign a consent form.

B.2 Stage 1 Consent Form

Queen Mary
University of London

Consent form

Please complete this form after you have read the Information Sheet and/or listened
to an explanation about the research.

Title of Study: Generating Affective 'Avant-Garde Animation' through Neurological and Physiological interaction relationships

Queen Mary Research Ethics Committee Ref: ________________

• Thank you for considering taking part in this research. The person organizing the research must explain the project to you before you agree to take part.

• If you have any questions arising from the Information Sheet or explanation already given to you, please ask the researcher before you decide whether to join in. You will be given a copy of this Consent Form to keep and refer to at any time.

• I understand that if I decide at any other time during the research that I no longer wish to participate in this project, I can notify the researchers involved and be withdrawn from it immediately.

• I consent to the processing of my personal information for the purposes of this research study. I understand that such information will be treated as strictly confidential and handled in accordance with the provisions of the Data Protection Act 1998.

• I consent to the use of images taken during the experiment to be used only in publications relating to this research (optional please tick the appropriate box, below)

☐ I agree to images of my interaction being used in publications related to this study.

☐ I do not agree to images of my interaction being used in publications related to this study.

Participant's Statement:

I ____ agree that the research project named above has been explained to me to my satisfaction and I agree to take part in the study. I have read both the notes written above and the Information Sheet about the project, and understand what the research study involves.

Signed: Date:

Investigator's Statement:

I ____ confirm that I have carefully explained the nature, demands and any foreseeable risks (where applicable) of the proposed research to the volunteer

B.3 Stage 1 SAM Test

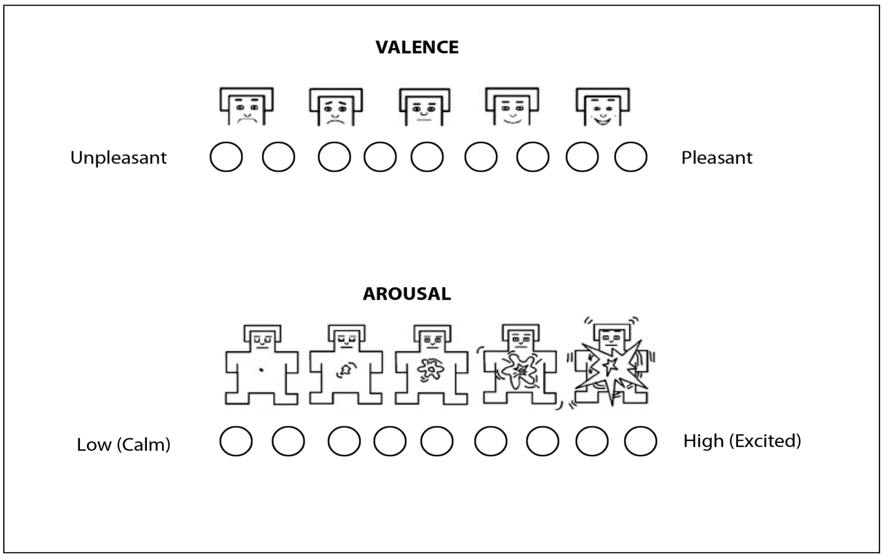
Queen Mary
University of London

Emotional Responses to Film & Video Clips (Part 1).

Name:

Film Clip Number:

Section A:



Section B:

(i) Please circle the emotion that most describes what you felt.

Afraid Amused Anxious Disgusted Joyful Sad Happy Content Neutral

Other………………..

(ii) Was the feeling consistent throughout the clip? Y / N (Please describe)

Section C:

(iii) Did you like the Film clip? Y / N (please circle)

(iv) What emotional feeling do you think this film clip might have tried to evoke? (please circle)

Afraid Amused Anxious Disgusted Joyful Sad Happy Content Neutral

Other………………..

(v) Have you seen the movie this clip was taken from? Y / N (please circle)

B.4 Stage 2 Information Sheet.

Queen Mary
University of London

Information sheet

Research study [title]: Generating Affective 'Avant-Garde Animation' through Neurological and Physiological interaction relationships (**QMREC2012/45**)

We would like to invite you without any obligation to be part of the research project described below. You should only agree to take part if you would like to. No disadvantages will arise should you decline to be involved.

Before making your decision, please read the following information carefully. This will inform you of the reasons for the research and inform you of what you will be asked to do if you decide to take part. Please do not hesitate to ask if there is anything that is not clear to you, or if you would like further information.

If you decide to take part, you will be asked to sign the attached form as proof of your agreement. Please note that you are still free to withdraw from the research study at any time and without giving a reason.

Details of study

The larger aims of this Research are to investigate the neural correlations of dynamic emotional states through Electroencephalographic (EEG) readings, in order to determine control measures for driving Affective Interactive Animations,

This particular experiment will take place in the Performance Lab at Queen Mary University, London. You will be guided to a quiet area, where the experimenter will fit an EEG wireless headset to your head, which measures real-time electrical activity in the brain. The experimenter will take a short baseline recording. Once complete you will be guided to a a designated desktop computer for the experiment.

A series of short film clips will be shown, and after each you will be prompted to fill in a short questionnaire. The experiment will last no more than 2 hours, If you do decide to take part you will be given this information sheet and be asked to sign a consent form.

B.5 Stage 2 Consent Sheet.

Queen Mary
University of London

Consent form

Please complete this form after you have read the Information Sheet and/or listened
to an explanation about the research.

Title of Study: Generating Affective 'Avant-Garde Animation' through Neurological and Physiological interaction relationships

Queen Mary Research Ethics Committee Ref: ________________

• Thank you for considering taking part in this research. The person organizing the research must explain the project to you before you agree to take part.

• If you have any questions arising from the Information Sheet or explanation already given to you, please ask the researcher before you decide whether to join in. You will be given a copy of this Consent Form to keep and refer to at any time.

• I understand that if I decide at any other time during the research that I no longer wish to participate in this project, I can notify the researchers involved and be withdrawn from it immediately.

• I consent to the processing of my personal information for the purposes of this research study. I understand that such information will be treated as strictly confidential and handled in accordance with the provisions of the Data Protection Act 1998.

• I consent to the use of images taken during the experiment to be used only in publications relating to this research (optional please tick the appropriate box, below)

☐ I agree to images of my interaction being used in publications related to this study.

☐ I do not agree to images of my interaction being used in publications related to this study.

Participant's Statement:

I ____ agree that the research project named above has been explained to me to my satisfaction and I agree to take part in the study. I have read both the notes written above and the Information Sheet about the project, and understand what the research study involves.

Signed: Date:

Investigator's Statement:

I ____ confirm that I have carefully explained the nature, demands and any foreseeable risks (where applicable) of the proposed research to the volunteer

B.6 Stage 2 Pre-Experiment Questionnaire.

Queen Mary
University of London

Pre-Experiment Questionnaire
Emotional Responses to Film & Video Clips (Part 2)

NAME

AGE

GENDER MALE / FEMALE * please cirle

CURRENT OCCUPATION

NATIONALITY

HANDEDNESS

HOW WOULD YOU DESCRIBE YOUR ETHIC BACKGROUND ? * please circle/

(a) Black or Black British
Caribbean
African
Any other Black background.

b) White
British
Irish
Any other White background

(c) Asian or Asian British
Indian
Pakistani
Bangladeshi
Any other Asian background within

(d) Mixed
White and Black Caribbean
White and Black African
White and Asian
Black and Asian
Any other Mixed

(e) Chinese or other Ethnic Group
Chinese
Any other Ethnic Group

(f) Rather Not Say
Rather not say.

B.7 Stage 2 SAM Test.

Queen Mary
University of London

Emotional Responses to Film & Video Clips (Part 2)

Name :

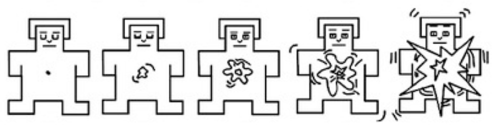
Film Clip Number :

SECTION A :

VALENCE

Unpleasant ○ ○ ○ ○ ○ ○ ○ ○ ○ Pleasant

AROUSAL



Low (Calm) ○ ○ ○ ○ ○ ○ ○ ○ ○ High (Excited)

SECTION B :

(i) Please circle the emotion that most describes what you felt.

Afraid Amused Anxious Disgusted Joyful Sad Happy Content Neutral

Other ....

(ii) Was the feeling consistent throughout the clip ? Y/N .

Bibliography

Aftanas, L. I., Reva, N. V., Varlamov, A. A., Pavlov, S. V., & Makhnev, V. P. (2004). Analysis of evoked EEG synchronization and desynchronization in conditions of emotional activation in humans: Temporal and topographic characteristics. *Neuroscience and Behavioral Physiology, 34,* (8). doi: 10.1023/B:NEAB.0000038139.39812.eb

AlZoubi, O., Calvo, R. A., & Stevens, R. H. (2009). Classification of EEG for affect recognition: An adaptive approach. *AI 2009: Advances in Artificial Intelligence, Lecture Notes in Computer Science, 58(*66) 52-61. doi : 10.1007/978-3-642-10439-8_6

Amabile, T. M. (1996). Creativity in context: Update to the social psychology of creativity. Boulder, CO: Westview Press.

Amabile, T. M. (2012). “Componential Theory of Creativity.” *Harvard Business School Working Paper #12-096.* http://dx.doi.org/10.4135/9781452276090

Badcock, N. A., Mousikou, P., Mahajan, Y., Lissa, P. D., Thie, J., & McArthur, G. (2013). Validation of the emotiv epoc EEG gaming system for measuring research quality auditory ERPs. *PeerJ 1:e38*. http://dx.doi.org/10.7717/peerj.38

Bard, P. (1928). "A diencephalic mechanism for the expression of rage with special reference to the sympathetic nervous system". *American Journal of Physiology 84*, 490–516.

Barrett, L. F. (2009). Variety is the spice of life: A psychological construction approach to understanding variability in emotion. *Cognition and Emotion, 23,* (7), 1284-1306. http://dx.doi.org/10.1080/02699930902985894

Brandao. P. (2012). Abstracting information on Body Area Networks. *Technical Report Number 812. University of Cambridge computer laboratory.* Retrieved from http://www.cl.cam.ac.uk/techreports/

Bos, D. O. (2006). EEG-based emotion recognition. The influence of visual and auditory stimuli. Retrieved from http://uais.lzu.edu.cn/uploads/soft/101215/15-101215200248.pdf

Cannon, W. B. (1927). The James-Lange theory of emotions: A critical examination and an alternative theory. *American Journal of Psychology;39:106–124.* http://dx.doi.org/10.2307/1415404

Carter, R. (2009). The brain book. Dorling Kindersley Limited. London.

Cavazza, M., Charles, F., Aranyo, G., Porteous, J., Gilroy, S., Ra, G., Keynan, N. J., Cohen, A., Jackont, G., Jacob, Y., Soreq, E., Klovatch, I., Hendler, T. (2014). Towards emotional regulation through neurofeedback. Proceedings from : *The 5th Augmented Human International Conference* (AH '14). March 7-9, 2014, Kobe, Japan. http://dx.doi.org/10.1145/2582051.2582093

Cavazza, M., Aranyi, G., Charles, F., Porteous, J., Gilroy, S., Jackont, G., Klovatch, I, Raz, G, Keynan, N. J, Cohen, A., Hendler, T. (2014). Frontal alpha asymmetry Neuro-feedback for brain-computer interfaces. *Proceedings of the 6th International Brain-Computer Interface Conference 2014, Graz, Austria, September 16-19.*

Cecotti, H. (2010). Spelling with brain-computer interfaces - Current trends and prospects. 5th *French Conference on Computational Neuroscience (Neurocomp 2010), Lyon : France.* http://dx.doi.org/10.1016/j.jphysparis.2011.08.003

Chanel, G., Kronegg, J., Grandjean, D., & Pun, T. (2005). Emotion assessment: Arousal evaluation using EEG's and peripheral physiological signals. Retrieved from http://vision.unige.ch/publications/postscript/2005/RAPPORT_TECH_05.02.pdf http://dx.doi.org/10.1007/11848035_70

Chanel, G , Kierkels, J. H. M., Soleymani, M., Pun, T. (2009). Short-term emotion assessment in a recall paradigm. *Int. J. Human-Computer Studies 67, 607–627.* http://dx.doi.org/10.1016/j.ijhcs.2009.03.005

Chen, M., Gonzalez, S., Vasilakos, A., Cao, H., & Leung, V. C. M. (2011). Body Area Networks : A survey. *Mobile Networks and Applications . April, 16,* (2), 171-193. http://dx.doi.org/10.1007/s11036-010-0260-8

Coan, J. A., & Allen, J. J. B. (2003). The state and trait nature of frontal EEG asymmetry in emotion. in Hugdahl, K., & Davidson, R. J. (Eds), The asymmetrical brain. Cambridge, MIT Press, 2003

Coan, J. A., & Allen, J. J. B. (2004). Frontal EEG asymmetry as a moderator and mediator of emotion. *Biological Psychology. No67*, 7–49. http://dx.doi.org/10.1016/j.biopsycho.2004.03.002

Costello, B., & Edmonds, E. (2007). A study in play, pleasure, and interaction design. *Proceedings of the 2007 conference on Designing pleasurable products and interface DPPI* 07. 76-91. http://dx.doi.org/10.1145/1314161.1314168

Damasio, A. R. (1999). The feeling of what happens: Body and emotion in the making of consciousness. William Heinemann. London

Damasio, A. R., Grabowski, T. J., Bechara, A., Damasio, H., Ponto, L. L. B., Parvizi, J., & Hichwa, R. D. (Oct 2000). Subcortical and cortical brain activity during the feeling of self-generated emotions. *Nature Neuroscience. 3(10).* 1049- 1056. http://dx.doi.org/10.1038/79871

Darwin, C. (1934). The expression of the emotions in man and animals. Watts & Co. London.

Davies, I, & Robinson, P. (2011). Emotional investment in naturalistic data collection. *ACII'11 Proceedings of the 4th international conference on Affective computing and intelligent interaction - Volume Part I.* 467-476. http://dx.doi.org/10.1007/978-3-642-24600-5_50

Davidson, R. J., & Fox, N. A. (17 Dec1982) Asymmetrical brain activity discriminates between positive and negative affective stimuli in human infants . *Science, New Series. 218(4578).* 1235- 1237. http://dx.doi.org/10.1126/science.7146906

Davidson, R. J. (1992). Anterior cerebral asymmetry and the nature of emotion. *Brain and Cognition. 20, 125-151.* http://dx.doi.org/10.1016/0278-2626(92)90065-t

Davidson, R. J. (1998). EEG measures of cerebral asymmetry: Conceptual and methodological issues. *International Journal of Neuroscience. 39.* 71-89. http://dx.doi.org/10.3109/00207458808985694

Davidson, R. J. (2012). The emotional life of your brain. Hodder & Soughton. London

Debener. S., Minow, F., Emkes, R., Gandras, K., & De Vos,. M. (2012). How about taking a low-cost, small, and wireless EEG for a walk? *Psychophysiology 49* . 1449–1453. http://dx.doi.org/10.1111/j.1469-8986.2012.01471.x

Elliot, A. J., & Covington, M.V. (2001). Approach and avoidance motivation. *Educational Psychology Review. 13,*(2) . http://dx.doi.org/10.4324/9780203888148.ch1

Ekman, P., & Friesen, W.V. (1975) Unmasking the face; A guide to recognizing emotions from facial expression. Prentice-Hall, inc. Englewood cliffs, New Jersey.

Ekman, P., Davidson, R. J., & Friesen, W. V. (1990). The duchenne smile: Emotional expression and brain physiology II. *Journal of Personality and Social Psychology. 58,* (2). 342-353. http://dx.doi.org/10.1037//0022-3514.58.2.342

Ertin, E., Raij, A., Stohs, N., Al'Absi, M., Kumar, S., & Shah, S. (2011). AutoSense : Unobtrusively wearable sensor suite for inferring the onset, causality, and consequences of stress in the field. *Proceedings of the 9th ACM SenSys,* Seattle, WA. http://dx.doi.org/10.1145/2070942.2071027

Foucher, J. R., Otzenberger, H., & Gounot, D. (2004). Where arousal meets attention: a simultaneous fMRI and EEG recording study. *NeuroImage 22* . 688– 697. http://dx.doi.org/10.1016/j.neuroimage.2004.01.048

Fox, N. A & Davidson, R. J. (1986). Taste-elicited changes in facial signs of emotion and the asymmetry of brain electrical activity in human newborns. *Neuropsychologia. 24. 417-422.* http://dx.doi.org/10.1016/0028-3932(86)90028-x

Fox, N. A., & Davidson, R. J. (1987). Electroencephalogram asymmetry in response to the approach of a stranger and maternal separation in 10 month-old infants. *Developmental Psychology. 23,*(2) 233-240. http://dx.doi.org/10.1037//0012-1649.23.2.233

Gendron, M, Barrett, L. F. (October 2009). Reconstructing the past: A century of ideas about emotion in psychology. *Emotion Review. 1; 1(4)*: 316–339. http://dx.doi.org/10.1177/1754073909338877

Gloor, P. (*Trans & Ed).* (1969). Hans Berger on the Electroencephalogram of man : The fourteen original reports on the human Electroencephalogram. Electroencephalography and Clinical Neurophysiology supplement no 28. Elsevier Publishing Company. Amsterdam. 1969

Goldman, A. (2004). Evaluating art. in Kivy (ed).The Blackwell Guide to Aesthetics. Blackwell Publishing. 93-108. http://dx.doi.org/10.1002/9780470756645.ch5

Goleman, D. (1996). Emotional intelligence: Why it can matter more than IQ. Bloomsbury. London

Griffiths, P. E, & Scarantino, A. Emotions in the wild: The situated perspective on emotion. in Robbins, P. and Aydede, M. (Eds) Cambridge handbook of situated cognition, Cambridge: Cambridge University Press, 437-453. (2009). http://dx.doi.org/10.1017/cbo9780511816826.023

Hagemann, D, Naumann, E., & Thayer, J.F. (2001). The quest for the EEG reference revisited; A glance from brain asymmetry research. *Psychophysiology. No 38*, 847-857. Cambridge University Press. http://dx.doi.org/10.1111/1469-8986.3850847

Harmon-Jones, E., & Allen, J. J. (1998). Anger and frontal brain activity : EEG asymmetry consistent with approach motivation despite negative affective valence. *Journal of Personality and Social Psychology .74*, (5).1310 – 1316. http://dx.doi.org/10.1037//0022-3514.74.5.1310

Huang, D., Guan, C., Ang, K. K., Zhang, H., & Pan, Y. (2012). Asymmetric spatial pattern for EEG-based emotion detection. *WCCI 2012 IEEE World Congress on Computational Intelligence June, 10-15, 2012 - Brisbane, Australia.* http://dx.doi.org/10.1109/ijcnn.2012.6252390

Healey, J. A. (2000). Wearable and automotive systems for affect recognition from physiology. (PhD thesis). Massachusetts Institute of Technology, Cambridge, MA, 2000

Healey, J. A. (2011) Recording affect in the field: towards methods and metrics for improving ground truth labels. *Affective Computing and Intelligent Interaction Lecture Notes in Computer Science Volume 6974,* 107-116. http://dx.doi.org/10.1007/978-3-642-24600-5_14

Höök. K. (2009). Affective loop experiences: designing for interactional embodiment *Phil.Trans. R. Soc. B 2009 364, 3585-3595.* http://dx.doi.org/10.1098/rstb.2009.0202

Horlings, R. Emotion recognition using brain activity. Man-Machine Interaction Group TU Delft. 27 March 2008. Retrieved from http p://www.kbs.twi.tudelft.nl/ Publications/ MSc/2008-Horlings-Msc.html

Isomursu, M.,Tahti, M., Vainamö, S., & Kuutti, K. (2007) Experimental evaluation of five methods for collecting emotions in field settings with mobile applications. *Int. J. Human-Computer Studies 65* . 404–418. http://dx.doi.org/10.1016/j.ijhcs.2006.11.007

Izard, C. (1991). The psychology of the emotions. Springer. Austria.

James, W. (1884). What is an emotion? In *Mind, 9,* (34), 188- 205

Jirayucharoensak, S., Pan-Ngum, S., & Israsena, P. (2014). EEG-Based emotion recognition using deep learning network with principal component based covariate shift adaptation. *The Scientific World Journal Volume 2014.* http://dx.doi.org/10.1155/2014/627892

Kagan, J. (2007). What is emotion ? : History, measures and meaning. Yale University Press.

Khalili, Z., & Moradi, M. H. (2008). Emotion detection using brain and peripheral Signals. *Biomedical Engineering Conference, 2008. CIBEC 2008. Cairo International* . http://dx.doi.org/10.1109/cibec.2008.4786096

Kim, D. H., Lu, N., Ghaffari, R., & Rogers, J. A. (2012). Inorganic semiconductor nanomaterials for flexible and stretchable bio-integrated electronics. *NPG Asia Materials (2012).* http://dx.doi.org/10.1038/am.2012.27

Koelstra, S., & Patras, I. (2013). Fusion of facial expressions and *EEG for implicit affective tagging. Image and Vision Computing. 31*, 164-174. http://dx.doi.org/10.1016/j.imavis.2012.10.002

Koelstra, S., Yazdani, A., Soleymani, M., M ̈uhl, C., Lee, J., Nijholt, A., Pun, T., Ebrahimi, T., & Patras, I. (2010). Single trial classification of EEG and peripheral physiological signals for recognition of emotions induced by music videos. *In: Proceedings 2010 International Conference on Brain Informatics (BI 2010), 28-30 August, Toronto.89-100.* http://dx.doi.org/10.1007/978-3-642-15314-3_9

Koelstra, S., M¨uhl, C., Soleymani, M., Lee, J-S., Yazdani, A., Ebrahimi, T., Pun, T., Nijholt, A., & Patras, I. (2012). DEAP: A database for emotion analysis using physiological signals. S*pecial Issue on Naturalistic Affect Resources for System Building and Evaluation of the IEEE Trans. on Affective Computing. Volume: 3* ,(1). 18-31. http://dx.doi.org/10.1109/t-affc.2011.15

Kroupi, S., Yazdani, A., & Ebrahimi, T. (2011). EEG correlates of different emotional states elicited during watching music videos. *ProceedingACII'11 Proceedings of the 4th international conference on Affective computing and intelligent interaction - Volume Part II* ,457-466 Springer-Verlag Berlin, Heidelberg. http://dx.doi.org/10.1007/978-3-642-24571-8_58

Kumar, P. S., Arumuganathan, R., Sivakumar, K., & Vimal, C. (December 2008). Removal of ocular artifacts in the EEG through wavelet transform without using an EOG reference channel. *Int. J. Open Problems Compt. Math.. 1*,(3). http://dx.doi.org/10.1109/icosp.2008.4697569

Lazarus, R.S. (1991). Emotion and adaptation. New York, NY, US: Oxford University Press.

Ledoux. J. (1996). The emotional brain : The mysterious underpinnings of emotional life. Simon & Schuster. New York.

Leonov, V., Yazicioglu, R.F., Merken, P., Van Hoof, C., Vullers, R. J. M., & Gyselinckx, B.(October 2008) Wearable autonomous wireless Electroencephalography system fully powered by human body heat . *Sensors, Oct 2008* .1269 – 1272. http://dx.doi.org/10.1109/icsens.2008.4716675

LeVan, P., Urrestarazu, E., & Gotman, J. (2006). A system for automatic artifact removal in ictal scalp EEG based on independent component analysis and Bayesian classification. *Clinical Neurophysiology 117* .912–927. http://dx.doi.org/10.1016/j.clinph.2005.12.013

Lindquist, K. A., Wager, T. D., Kober, H., Bliss-Moreau, E., & Barrett, L.F. (2012). The brain basis of emotion: A meta-analytic review. *Behavioural and Brain Science 35,* 121–202. http://dx.doi.org/10.1017/s0140525x11000446

Liu, Y., & Sourina, O. (2012). EEG-based dominance level recognition for emotion-enabled Interaction. *2012 IEEE International Conference on Multimedia and Expo.* http://dx.doi.org/10.1109/icme.2012.20

Liu, Y., Sourina, O., & Hafiyyandi, M. R. (2013) EEG-based emotion-adaptive advertising. *2013 Humaine Association Conference on Affective Computing and Intelligent Interaction.* http://dx.doi.org/10.1109/acii.2013.158

MacLean, P. (1990) The triune brain in evolution. Plenum Press. New York.

McKeown, G., Valstar, M., Cowie, R., Pantic, M., Schroder, M. (2012) The SEMAINE database: Annotated multimodal records of emotionally colored conversations between a person and a limited agent. *IEEE Transactions on Affective Computing. 3*,(1) . http://dx.doi.org/10.1109/t-affc.2011.20

Mehrabain, A., and Russell, J. A. (1974). An approach to environmental psychology. The MIT Press. Cambridge, MA.

Mikhail, M., El-Ayat, K., Coan, J. A., & Allen, J. B. (2013). Using minimal number of electrodes for emotion detection using brain signals produced from a new elicitation technique. *Int. J. Autonomous and Adaptive Communications Systems. 6,* (1). http://dx.doi.org/10.1504/ijaacs.2013.050696

Ormerod, W. (2006). Richard Caton (1842-1926) : Pioneer electrophysiologist and cardiologist. *Journal of Medical Biography .14* .30-35. http://dx.doi.org/10.1258/j.jmb.2006.04-22

Ortony, A., Collins, GL., & Clore, A. (1988). The cognitive structure of emotions. Cambridge University Press. Cambridge.

Panksepp, J. (1998). Affective Neuroscience : the foundations of human and animal emotions .Oxford university press . Oxford 1998.

Panksepp, J. (2000). Emotions as natural kinds within the mammalian brain. 137-156 In Handbook of emotions. Second edition. Ed. Michael Lewis, J.M Haviland-Jones. The Guildford press. New York.

Panksepp, J. Interview (2012). Brain Science podcast with Ginger Campbell,MD episode 65. Interview with Dr Jaak Panksepp, authour of affective neurocsience: The foundations of human and animal emotions. Aired January 13, 2010. Retrieved from: http://brainsciencepodcast.com/bsp/2010/1/13/affective-neuroscience-with-jaak-panksepp-bsp-65.html

Papez, J. W. (1937). A proposed mechanism of emotion. *Archives of Neurology and Psychiatry.* 725-743. http://dx.doi.org/10.1001/archneurpsyc.1937.02260220069003

Peter, C., Ebert, E., Beikirch, H. (2005). A wearable multi-sensor system for mobile acquisition of emotion-related physiological data. *ACII 2005, LNCS 3784,* . 691 – 698. http://dx.doi.org/10.1007/11573548_89

Picard, R. W. (1997). Affective computing. The MIT Press. Cambridge. MA

Picard, R. W., & Scheirer, J. (2001) The Galvactivator: a glove that senses and communicates skin conductivity. *Proc. Ninth Int. Conf. on Human– Computer Interaction, New Orleans, LA* 1538-1542. Retrieved from: http://vismod.media.mit.edu/pub/tech-reports/TR-542.pdf

Picard, R. W., Liberate yourself from the lab: Q Sensor measures EDA in the wild. Affective Q white paper. Retrieved from: http://www-assets.affectiva.com/assets/Whitepaper_Q_Measu-ring_EDA_in_the_Wild.pdf

Picard, R. W. (2010) Emotion research by the people, for the People. *Emotion Review 2010 2,* (3), 250-4. http://dx.doi.org/10.1177/1754073910364256

Plutchik, R. (2001). The nature of emotions. *American Scientist, 89.* 344-350. http://dx.doi.org/10.1511/2001.28.739

Plutchik, R. (2003). Emotions and life; Perspectives from psychology , biology, and evolution. American Psychological Association. Washington. Dc.

Poh, M., Swenson, N. C., & Picard, R. W. (May 2010). A wearable sensor for unobtrusive, long-term assessment of electrodermal activity. *IEEE Transactions on Biomedical Engineering.57,* (5). http://dx.doi.org/10.1109/tbme.2009.2038487

Rolls, E.T. (2000). Précis of the brain and emotion. B*ehavioural and Brain Sciences. 23,* (2*)*, 177-234. http://dx.doi.org/10.1017/s0140525x00002429

Russ, S. W. (1993). Affect & creativity : The role of affect & play in the creative. Routledge. London.

Russell, J. A. (1980). A circumplex model of affect. *Journal of Personality and Social Psychology. 39,* (6),.1161-1178. http://dx.doi.org/10.1037/h0077714

Russell, J. A. (1994). Is there universal recognition of emotion from facial expression ? A review of the cross-cultural studies. *Psychological Bulletin. 115*,(1), 102-141. http://dx.doi.org/10.1037/0033-2909.115.1.102

Russel, J. A. (2009). Emotion, core affect, and psychological construction. *Cognition and Emotion. 23*, (7), 1259-1283. http://dx.doi.org/10.1080/02699930902809375

Sanei, S., & Chambers, J. A. (2007). EEG signal processing. John Wiley & Sons ltd. Chicester. Uk.

Salovey, P., & Mayer, J. (1990). Emotional intelligence*. Imagination, Cognition and Personality, 9*, 185-211. Retrieved from: http://www.unh.edu/emotional_intelligence/EIAssets/EmotionalIntelligenceProper/EI1990%20Emotional%20Intelligence.pdf

Schachter, S., & Singer, J. (1962). Cognitive, social and physiological determinants of emotional state. *Psychological review. 69*, (5), 379 -399. http://dx.doi.org/10.1037/h0046234

Scherers, K. (1984 a). On the nature and function of emotion: A component process approach. In Scherer, K.R., & Ekman, P. Approaches to Emotion. Lawrence Erlbaum. Hillsdale, N. J.

Scherer. K. (1984). Emotion as a multicomponent process. A model and some cross cultural data. *in P.Shaver (ed) review of personality and social Psychology, Vol5, Beverly Hills, CA; Sage (b),*37-63

Scholsberg, H. (1954). Three dimensions of emotion. *Psychological Review. Vol. 61,* (2)*,* 81-88. http://dx.doi.org/10.1037/h0054570

Sternberg, R. J., & Lubart, T. I. (1996). Investing in creativity. *American Psychologist 51,* (7), 677–688. http://dx.doi.org/10.1037//0003-066x.51.7.677

Sternberg, R. J. (2005). Creativity or creative's?. *Int. J. Human-Computer Studies 63*, (4-5),370–382. http://dx.doi.org/10.1016/j.ijhcs.2005.04.003

Swartz, B. E., & Goldensohn, E.S. (1998). Timeline of the history of EEG and associated fields. *Electroencephalography and Clinical Neurophysiology, 106,* 173-176

Soleymani, M., Pantic, M., & Pun, T. (April-June 2012). Multimodal emotion recognition. *In IEEE Transactions on Affective Computing. 3,*(2). http://dx.doi.org/10.1109/t-affc.2011.37

Soleymani, M., Lichtenauer, J., Pun, T., & Pantic, M. (2012). A Multimodal Database for Affect Recognition and Implicit Tagging. *IEEE Transactions on Affective Computing. 3,* (1). http://dx.doi.org/10.1109/t-affc.2011.25

Sutton, S. K., & Davidson, R. J. (2000). Prefrontal brain electrical asymmetry predicts the evaluation of affective stimuli. *Neuropsychologia, No 38*, (17*),* 1723 – 1733. http://dx.doi.org/10.1016/s0028-3932(00)00076-2

Teplan, M. (2002) Fundamentals of EEG measurement. *Measurement Science Review. 2,* Section 2.

Tomarken, A. J., Davidson, R. J., & Henriques, J. B. (1990). Resting frontal brain asymmetry predicts affective responses to films. *Journal of Personality and Social Psychology . 59,(* 4), 791-801. http://dx.doi.org/10.1037//0022-3514.59.4.791

Tomkins, S. S. (1982). Affect theory in emotion in the human face. Second Edition Edited by Ekman. P., Cambridge University Press. Cambridge .

Viola, F. C., Thorne, J., Edmonds, B., Schneider, T., Eichele, T., & Debener, S. (2009). Semi-automatic identification of independent components representing EEG artifact. *In Clinical Neurophysiology 120*, (5), 868–877. http://dx.doi.org/10.1016/j.clinph.2009.01.015

Wheeler, R. E., Davidson, R. J., & Tomarken, A. J. (1993). Frontal brain asymmetry and emotional reactivity : A biological substrate of affective style. In P*yschophysiology,30,* (1), 82-89. http://dx.doi.org/10.1111/j.1469-8986.1993.tb03207.x

Winkler, I., Haufe, S., & Tangermann, M. (2011). Automatic classification of artifactual ICA components for artefact removal in EEG Signals. *Behavioural and Brain Functions*, 7, (1), 30. http://dx.doi.org/10.1186/1744-9081-7-30

Woodman, G.F. (2010). A brief introduction to the use of event-related potentials in studies of perception and attention. *Attention, Perception, & Psychophysics .72* (8), 2031-2046. http://dx.doi.org/10.3758/bf03196680

Xiaowei, L., Qinglin, Z., Li, L., Hong, P., Yanbing, Q., Chengsheng, M., Zheng, M., & Quanying, L. (2010). Improve Affective Learning with EEG Approach. *Computing and Informatics, 29*, 557–570

Yeung, R. (2006). The Rules of EQ. Marshall Cavendish Buisness. Singapore.

Yoon, H. J., & Chung, S. Y. (2011). EEG Spectral Analysis in Valence and Arousal Dimensions of Emotion. *11th International Conference on Control, Automation and Systems. KINTEX,* Gyeonggi-do, Korea. October 2011, 1319-1322.

Zeki, S. (1999). Inner Vision : An Exploration of Art and the Brain. Oxford University Press. Oxford